METALS IN CHEMISTRY AND BIOLOGY: COMPUTATIONAL CHEMISTRY STUDIES

Adriana Dinescu, B.S., M.S.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2007

APPROVED:

Thomas R. Cundari, Major Professor
Mohammad A. Omary, Committee Member
Martin Schwartz, Committee Member
Angela K. Wilson, Committee Member
Ruthanne D. Thomas, Chair of the Department of Chemistry
Sandra L. Terrell, Dean of the Robert B. Toulouse School of Graduate Studies

Dinescu, Adriana. Metals in Chemistry and Biology: Computational Chemistry Studies. Doctor of Philosophy (Inorganic Chemistry), May 2007, 86 pp., 10 tables, 20 illustrations, references, 101 titles.

Numerous enzymatic reactions are controlled by the chemistry of metallic ions. This dissertation investigates the electronic properties of three transition metal (copper, chromium, and nickel) complexes and describes modeling studies performed on glutathione synthetase. (1) Copper nitrene complexes were computationally characterized, as these complexes have yet to be experimentally isolated. (2) Multireference calculations were carried out on a symmetric C_{2v} chromium dimer derived from the crystal structure of the $[(^{t}Bu_3SiO)Cr(\mu\text{-}OSi^{t}Bu_3)]_2$ complex. (3) The T-shaped geometry of a three-coordinate β-diketiminate nickel(I) complex with a CO ligand was compared and contrasted with isoelectronic and isosteric copper(II) complexes. (4) Glutathione synthetase (GS), an enzyme that belongs to the ATP-grasp superfamily, catalyzes the (Mg, ATP)-dependent biosynthesis of glutathione (GSH) from γ–glutamylcysteine and glycine. The free and reactant forms of human GS (wild-type and glycine mutants) were modeled computationally by employing molecular dynamics simulations, as these currently have not been structurally characterized.

ACKNOWLEDGEMENTS

I express my special thanks to my advisor, Professor Thomas R. Cundari, who guided me throughout my graduate studies at the University of North Texas. I would like to thank to collaborators Dr. Peter T. Wolczanski (Cornell University), Dr. Patrick L. Holland (University of Rochester), and Dr. Mary E. Anderson (Texas Woman's University) for their experimental contributions to this work. I would like to express my gratitude to Mary E. Anderson for her constant support during my graduate studies. The Anderson research group is gratefully acknowledged as well for providing the necessary data from experimental results.

I also want to acknowledge the support of the Chemical Computing Group for generously offering the MOE software. Support for this work has been provided by the University of North Texas, U.S. Department of Education (CASCaM), U.S. Department of Energy, and National Science Foundation.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF ILLUSTRATIONS

Page

CHAPTER 1

INTRODUCTION

Metal catalysts and metal catalysis are a foundation of the modern chemical industry. Although many industrial catalytic processes have been developed in the past, the chemical industry continually seeks new catalytic technologies that are more “green”, e.g., homogeneous catalysis or catalysts that use "green" metals such as copper. Since homogenous catalysts, especially those that incorporate transition metals, typically function effectively under milder conditions than heterogeneous catalysts, transition metals have become extremely important to the field of catalysis. The main reasons why transition metals are the essential constituent of many catalytic systems include their ability to stabilize widely disparate geometries, variable coordination numbers, a variety of ligand types, multiple oxidation states, and different spin states. A major advantage of homogeneous catalysis is selectivity – the ability to produce pure products in high yield. Much of the selectivity observed with soluble catalysts arises from the process control that is possible in the liquid phase.

Given the importance in catalysis of metastable chemical species that form in very low concentrations, and transition states for which direct experimental data is often very limited, designing efficient catalysts is one of the areas where computational chemistry can have a substantial impact in the chemical industry. In fact, computational chemistry has advanced so significantly that it has become an invaluable adjunct to experiment. The expansion of computer power and software packages for chemical research has opened the possibility for new computational investigations into the

structure, reaction mechanisms, and catalysis of organometallic complexes and metalloenzymes.

Currently, catalysis and the design of catalysts are of great practical importance as natural enzymes are generally considered to be among the most proficient catalysts. Numerous enzymatic reactions are controlled by the chemistry of metallic ions. While Mg(II) is the most frequent non-transition metal in the active site of metalloenzymes, several first-row transition metals (*e.g.,* Fe, Zn, Mn, Co, Cu, Ni, and V) also play key roles. This dissertation consists of several computational studies of chemical systems that incorporate Cu, Cr, Ni, and Mg metals.

Chapter 3 describes a computational investigation of copper nitrene complexes. Given that copper is a familiar redox metal serving a biological function, copper enzymes are intriguing catalysts for chemical reactions. Copper sites are found in oxygen carriers, mono-oxygenases, dioxygenases and oxidases.[1] Production of O_2 by photosynthetic organisms leads to oxidation of Cu(I) to Cu(II). It has been shown that in galactose oxidases the Cu(II) site is coupled to a substituted tyrosyl radical, while in amine oxidases, the Cu(II) site is associated with a quinone.[2] Therefore, study of copper complexes is useful for better understanding the reaction mechanisms that involve copper metalloenzymes and copper catalysts.

Eight Ni enzymes are known: urease, hydrogenase, CO-dehydrogenase (CODH), acetyl-coenzyme A synthase, methyl-coenzyme M reductase, Ni-superoxide dismutase, glyoxalase I enzymes, and cis-trans isomerase.[3] The crystal structure of acetyl-

coenzyme A synthase/CO dehydrogenase from *Moorella thermoacetica* shows that the protein contains two Ni-containing active sites that catalyze distinct CO transformations.[4] One of these active sites, the A-cluster, contains a Ni–Ni–[Fe_4S_4] species in which the central bridging Ni ion is square planar. It has been proposed that the bridging (proximal) Ni center binds CO and a methyl group, which then inserts into the Ni–CH_3 bond to form a Ni-bound acetyl group. Then, the Ni-bound acetyl group is attacked by CoA to yield the product.[5] The distal Ni center and the [Fe_4S_4] moiety are suggested to be necessary for tuning the activity of the proximal Ni atom through cysteinyl ligation around this bridging Ni centre.[6] In Chapter 5 the preferred geometry of a three coordinate nickel(I)-carbonyl complex is discussed.

Adenosine 5'-triphosphate (ATP), the universal energy carrier in the living cell, consists of adenosine linked to three phosphate groups. Removal of the γ-phosphate to form adenosine diphosphate (ADP) provides energy for use in other biological processes. Protein motors such as myosin, kinesin, ATP synthase, and various ATPases, use ATP as the energy source.[7] The use of ATP is "the most prevalent chemical reaction in the human body".[8] Mg(II) ions have a significant role in promoting ATP hydrolysis. Interaction between the Mg(II) ion and the γ-phosphate of ATP results in stabilization of the phosphate's negative charge in the transition state, which is thought to accelerate the rate of ATP hydrolysis.[9] Research on protein kinase and ATP-ase[10-11] has shown that ATP is bound with two Mg^{2+} ions, a "catalytic" Mg^{2+} bound to the γ-phosphate of ATP and a "structural" Mg^{2+} bound to the α and β-phosphates of ATP. When the substrate and Mg/ATP are bound to the enzyme, catalysis occurs rapidly and

product is formed. It is not surprising that ATP enzymes have been studied extensively; chapters 6 and 7 include the study of the reaction mechanism of glutathione biosynthesis, which involves ATP and Mg^{2+} as cofactors for the enzyme glutathione synthetase.

Experimental approaches are often insufficient to characterize fully the complex nature of enzymes or organometallic systems, and thereby, computational chemistry is essential for the assessment of electronic properties of the metal site, its reactivity, and the role of ligands or protein residues that influence the metal chemistry. Due to the great versatility of transition metal ions and the large size of biological systems, modeling this type of chemical systems represents a significant challenge to computational chemists. Characterization of the microscopic mechanisms by which enzymes carry out their designed catalytic role is a focal point in modern enzymology. Obtaining molecular level information is critical in predicting the behavior of enzymes and is essential for the design of new enzymes or modification of existing enzymes for specific catalytic tasks. In this research, modern computational approaches are used to calculate the properties of chemical and biological systems, which are expected to have future importance in catalysis and biocatalysis. The computational research presented herein complements efforts in experimental chemistry and is expected to promote better understanding of complex chemical reactions involving metal ions.

CHAPTER 2

GENERAL COMPUTATIONAL METHODS

2.1. Quantum Mechanics

Over the last decade, significant progress has been made in applying quantum mechanical methods to chemical problems involving the structures and reactions of molecules. Quantum mechanics (QM) states that the energy and other related properties may be calculated by solving the Schrödinger equation. However, the exact solutions to the Schrödinger equation are not computationally feasible for most molecules. Therefore, several families of electronic structure methods have been developed: semiempirical, *ab initio*, and density functional theory (DFT).

The most common type of *ab initio* method is the Hartree-Fock (HF) method[12] in which the primary approximation is the central field approximation. This approximation gives the average effect of the electron-electron repulsion instead of the explicit repulsion interaction. Due to the central field approximation, HF calculations do not include electron correlation and the energies resulting from these calculations are always greater than the exact energy.

Transition metal systems typically require explicit electron correlation and therefore HF is typically not suitable for these systems. Since post-HF *ab initio* methods and DFT include electron correlation, these methods are more appropriate for transition metals. One of the post-HF methods that include electron correlation is multiconfiguration self-consistent field (MCSCF),[13] which uses multiple determinants for

describing the wave function. MCSCF calculations are employed for near degenerate states and open shell systems, and can be very accurate, but the cost in CPU time is very high. In an MCSCF calculation not only the coefficients of the multiple determinants, but also the orbitals are optimized. This type of calculation requires the specification of the active orbitals to be included in the active space. An MCSCF calculation in which all combinations of the active space orbitals are included is called a complete active space self-consistent field (CASSCF)[14] calculation. CASSCF calculations give the maximum electron correlation in the valence region.

The DFT family of methods[15] has become very popular over the last decade because they are less computationally expensive when compared to post-HF methods. In practice, a DFT calculation involves a similar computational effort to that required for a HF calculation. The basis of DFT is that the energy of a molecule can be determined from the electron density instead of a wave function. The B3LYP hybrid functional[16] is the most widely used functional due to the good results obtained for a large range of compounds. This hybrid functional involves the Becke three term exchange functional (three fitting parameters are used) coupled with Lee, Yang, and Parr correlation functional.

For very large molecules such as enzymes with thousands of atoms, electronic methods are not feasible even on the largest computers. However, it is possible to model large molecules using molecular mechanics or a hybrid approach that involves both quantum mechanics and molecular mechanics (QM/MM).

2.2. Molecular Mechanics

The molecular mechanics (MM) energy expression consists of the sum of simple classical equations, which describe various parameters of the molecule: bond stretching, angle bending, torsions, electrostatic interactions, van der Waals forces, hydrogen bonding, etc. The set of MM equations with their associated constants is called a force field.[17] Force fields differ in the number of terms in the energy expression, the complexity of those terms, and the way in which the constants are determined. Since electrons are not explicitly included, electronic processes cannot be readily modeled with MM. The Assisted Model Building with Energy Refinement (AMBER)[18] force field was parameterized specifically for proteins and nucleic acids. AMBER uses five bonding and non-bonding terms along with a complex electrostatic treatment.

Molecular dynamics (MD)[19] is very powerful tool for conformational searches and involves simulation of the time dependent behavior of molecules by using molecular mechanics calculations. MD simulations are used to estimate equilibrium and dynamic properties of large systems; these properties express how atoms stretch, vibrate, and rotate about the bonds. The resultant potential energy for these systems is highly approximate due to representation of solvent, polarization effects, and electrostatic interactions. Minimization of this approximate energy function yields information on favorable (i.e., low energy) regions in conformational space.

2.3. Hybrid QM/MM Method

The QM/MM technique is used to model large systems for which full QM calculations are not feasible. A large system is computed by both a small model (typically, the metal and ligating atoms) and a complete system. The complete system is computed using only the lower level of theory, which is MM in this research. The model system is computed with both levels of theory (QM and MM). By combining these calculations, the energy for the complete system is obtained with the following equation:

$$E = E_{low,complete} + E_{high,model} - E_{low,model}$$

CHAPTER 3

BONDING AND STRUCTURE OF COPPER-NITRENES

3.1. Introduction

Copper catalysts are widely used for nitrene (NR, where R is most often a hydrocarbyl group) transfer, most notably the aziridination of unsaturated organic substrates such as olefins.[20] Several researchers have also demonstrated the feasibility of arylisocyanate synthesis by the reaction of CO with structurally characterized aryl-imido/nitrene complexes of the late transition metals. Hillhouse and coworkers have reacted a *bis*(phosphine)Ni-nitrene with CO to form arylisocyanates.[21] Bart *et al.*[22] and Peters et al.[23] have reported nitrene group transfer from structurally characterized late transition metal-nitrene complexes (*i.e.*, cobalt and iron) to CO to yield arylisocyanates (ArNCO).

While aziridination of olefins has been widely studied from an experimental point of view,[20] another attractive synthetic target is nitrene transfer to CO to yield isocyanates, which are widely used as intermediates for production of poly-isocyanates and polyurethanes.[24] There has been considerable interest in phosgene-free processes for the production of isocyanates from nitroaromatics.[24,25] Moiseev *et al.* reported the production of phenyl isocyanate from nitrobenzene with a Pd cluster catalyst; nitrenes were implicated as intermediates in this transformation.[26] Tsumura *et al.* filed a patent for the production of arylisocyanates from nitroaromatics by their reaction with CO.[27] Liu *et al.* discuss the formation of arylisocyanates from alkyl-nitro complexes by a copper zeolite catalyst.[28] Kober *et al.* report the formation of halogenated

arylisocyanates from halogenated nitroaromatics using a catalyst that is 5% $PdCl_2$ and 5% $CuCl_2$ supported on SiC.[29]

Although an isolated copper nitrene complex has not been reported, there are structurally characterized copper(I) amido complexes, which could act as precursors for a copper nitrene intermediate. Most of them contain chelating amide ligands[30] or, adopt dinuclear,[31] trinuclear,[32] or tetranuclear[32,33] structures with bridging amide groups, the nuclearity being determined by steric effects. Recently, a monomeric Cu(I) anilido system with a bulky bisphosphine ligand has been successfully synthesized and characterized;[34] its superior reactivity renders the copper amido complexes as potential effective catalysts. Warren and coworkers isolated a nitrene bound between two copper β-diketiminate groups.[35]

This work represents an analysis of the bonding, structure and reactivity of terminal copper-nitrenes ($L_nCu(NR)$) as such species have intrigued synthetic chemists in terms of their putative intermediacy in nitrene transfer catalysis.[20]

3.2. Computational Methods

B3LYP[16] geometry optimization utilized the Gaussian03[36] suite of programs. CASSCF[14] calculations were carried out with the GAMESS package.[37] QM/MM calculations employed the SIMOMM scheme,[38] integrated in GAMESS with the TINKER module.[39] The QM part was treated with CASSCF approach, while the MM part involved the MM3 force field.[40] Stevens' effective core potentials (ECPs) and valence basis sets

(VBSs) were employed.[41] Main group basis sets were augmented with d polarization functions.

Vibrational frequencies were calculated at all DFT optimized stationary points to confirm them as minima or transition states. Modeling of triplet species with density functional theory employed unrestricted Kohn-Sham methods.

3.3. Results and Discussion

Copper-nitrene complexes are proposed as key intermediates in copper-catalyzed nitrene transfer. Furthermore, isolated nitrenes of late transition metals such as Fe,[22,23] Co[23] and Ni[21] are known to transfer NR to CO to yield isocyanates.

Given their importance and experimental scarcity, a detailed investigation of the bonding in copper-nitrene complexes was undertaken. Models with phenylnitrene (NPh) as a ligand were studied. Borden and coworkers[42] have published extensively on the bonding and reactivity of NPh. In this research β-diketiminate ligands were chosen as a supporting ligand system with a phenylnitrene ligand. This choice was motivated by recent work from the Warren,[43] Gunnoe[44] and Holland[45] groups. The copper-containing portion (L'Cu, where L' is the parent β-diketiminate anion, $C_3N_2H_5^-$) of the nitrene models is formally a closed-shell d^{10}-Cu(I) ion and hence a singlet ground state is expected for L'Cu (a supposition supported by DFT calculations). For phenylnitrene (PhN) a singlet($^1A_2$)-triplet($^3A_2$) splitting of ~18 kcal/mol is obtained from a combination of multiconfiguration calculations and experiment.[42] The ground state of PhN is thus a triplet. Note that both of the aforementioned states are "open-shell" (*i.e.*, $\phi_1(\alpha)\phi_2(\beta)$

species, even the singlet, $^1A_2$-PhN. There is a "closed-shell" (*i.e.*, $\phi_1(\alpha)\phi_1(\beta)$) $^1A_1$ state of PhN that is calculated to be ~40 kcal/mol above the $^3A_2$ ground state.[42]

Given the bonding in its L'Cu and NPh constituents, one must address whether copper-nitrenes are more stable in a singlet or triplet state, and compare open- and closed-shell descriptions of the singlet states. The copper-nitrene ground spin state will depend on the nature and strength of the copper-nitrene bond. While a copper(III)-imido (*i.e.*, NR^{2-}) description is inconsistent with the group transfer chemistry of copper aziridination catalysts, any degree of multiple bonding in the Cu(NR) linkage is expected to shift the energetic balance away from a triplet and towards a singlet ground state.[46] Conversely, a triplet copper-nitrene ground state is expected to result from a weaker, less covalent copper-nitrene bond.

3.3.1. DFT Calculations

A DFT and CASSCF analysis was conducted for (β-diketiminate)Cu(NPh). The latter method, although computationally much more expensive than DFT, was deemed prudent in light of the results by Borden *et al.* on NPh.[42]

3.3.1.1. Triplet State

B3LYP/CEP-31G(d) geometry optimization of ground state triplet L'Cu(NPh) yields a linear nitrene coordination mode, which is coplanar with the β-diketiminate ligand ($Cu\text{-}N_{nitrene}$ = 1.779 Å; $N_{nitrene}\text{-}C_{ipso}$ = 1.331 Å; $Cu\text{-}N_{nitrene}\text{-}C_{ipso}$ = 180.0°). Unrestricted DFT computations on ^{3}L'CuNPh indicate minimal spin density on the β-diketiminate

supporting ligand. However, spin polarization effects are substantial as can be seen from large negative spin densities on the ipso and meta carbons of the phenyl substituent, Figure 3.1, in addition to positive spin density on the ortho and para carbons. Interactions of π type between nitrene nitrogens and aryl substituents are also reflected in the short $N_{nitrene}$-C_{ipso} distances.[42] The delocalization suggests the potential for substantial control of copper-nitrene reactivity through the choice of functional groups on the aryl substituent.

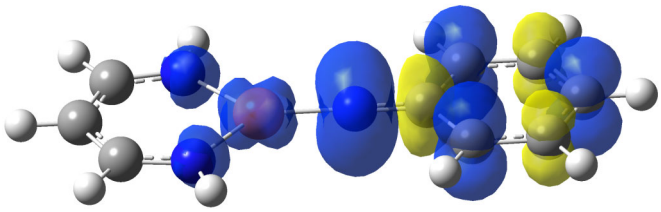


Figure 3.1. Spin density of (β-diketiminate)Cu(NPh). Blue – positive spin density; yellow – negative spin density.

3.3.1.2. Singlet State

A singlet state for L'Cu(NPh) was also geometry optimized by DFT methods and yielded a bent geometry (Cu-$N_{nitrene}$ = 1.726 Å; $N_{nitrene}$-C_{ipso} = 1.356 Å; Cu-$N_{nitrene}$-C_{ipso} = 147.8°). For L'Cu(NPh), the singlet is ~C_s symmetry. Unlike the triplet, the singlet copper-nitrene shows some evidence of copper-nitrogen multiple bonding (Figure 3.2, bottom). Like the triplet, there is delocalization of the nitrene pπ orbital onto the aryl substituent (Figure 3.2, middle) in the lowest energy singlet state of L'Cu(NPh).

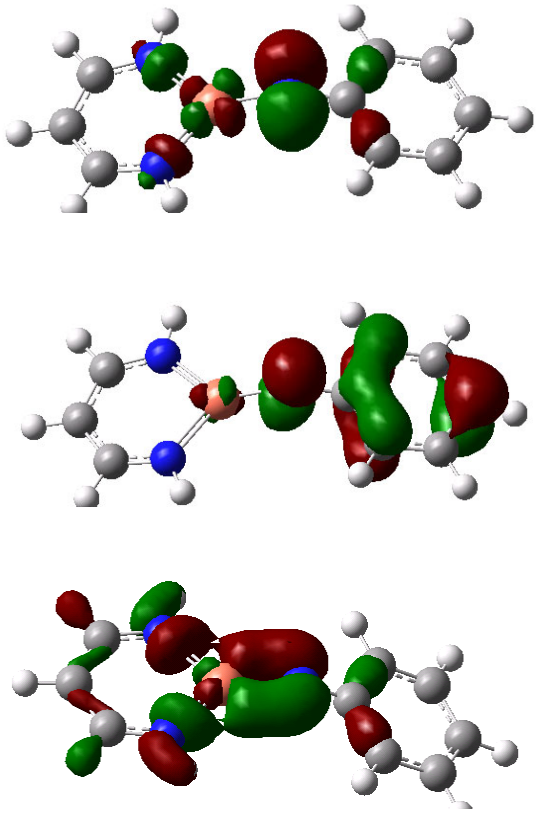


Figure 3.2. Frontier orbitals of singlet (β-diketiminate)Cu(NPh): LUKSO (top), HOKSO (middle), and HOKSO-3 (bottom). HOKSO = highest occupied Kohn-Sham orbital; LUKSO = lowest unoccupied Kohn-Sham orbital.

3.3.2. CASSCF Calculations

3.3.2.1. C_{2v} Symmetry

The CASSCF and DFT calculations yield different viewpoints of Cu-nitrene electronic structure, in particular singlet states. CASSCF/CEP-31G(d) geometry optimizations with a 10-orbital/10-electron active space were first carried out under C_{2v} symmetry for singlet and triplet (β-diketiminate)Cu(NPh) for all possible state symmetries. Several CASSCF active space sizes were investigated and the results were found to be similar, *i.e.*, the frontier b_1 and b_2 π orbitals of the copper-nitrene linkage

are the most strongly correlated orbitals. Hence, the discussion focuses on results from the largest CASSCF active space studied. The 10-orbital, 10-electron active space was chosen to incorporate electron correlation effects from the four highest energy doubly occupied orbitals that subtend the $a_1 + a_2 + b_1 + b_2$ representations, and the four lowest energy unoccupied orbitals that span the same representations, plus the singly occupied b_1 and b_2 of the Hartree-Fock reference wavefunction. The CASSCF(10,10) calculations yield a $^3A_2$ ground state within the constraint of C_{2v} symmetry. Low energy $^1A_1$, $^1A_2$ and $^3A_2$ states were also obtained and analyzed. Singlet and triplet states of B_1 and B_2 symmetry were calculated to be much higher in energy (>73 kcal/mol).

The $^3A_2$ state obtained by CASSCF geometry optimization of L'Cu(NPh) is consistent with the ground state triplet found by DFT calculations, since the two singly occupied Kohn-Sham orbitals subtend the b_1 and b_2 irreducible representations in the C_{2v} point group: $b_1 \otimes b_2 = a_2$. Furthermore, Borden *et al.* have concluded from experimental and computational data that the ground state of PhN is $^3A_2$, which is ~18 – 20 kcal/mol below the lowest energy singlet state ($^1A_2$) of NPh.[42] Hence, the of $^3A_2$-(β-diketiminate)Cu(NPh) state is conceptually the product of $^3A_2$-NPh and $^1A_1$-(β-diketiminate)Cu^I.

The singlet($^1A_2$)-triplet($^3A_2$) splitting at the CASSCF level of theory is considerably less than indicated by DFT calculations, *i.e.*, 2 kcal/mol, Figure 3.3 (left hand side). The lowest singlet $^1A_1$ is now only 5 kcal/mol higher in energy than the $^3A_2$ state of (β-diketiminate)Cu(NPh) for CASSCF/CEP-31G(d) geometries optimized within the constraint of C_{2v} symmetry. Hence, CASSCF calculations greatly stabilize both the open-

shell ($^1A_2$) and closed-shell ($^1A_1$) singlet states of the copper-nitrene. Pertinent structural data for C_{2v} nitrene complexes is summarized in Table 3.1.

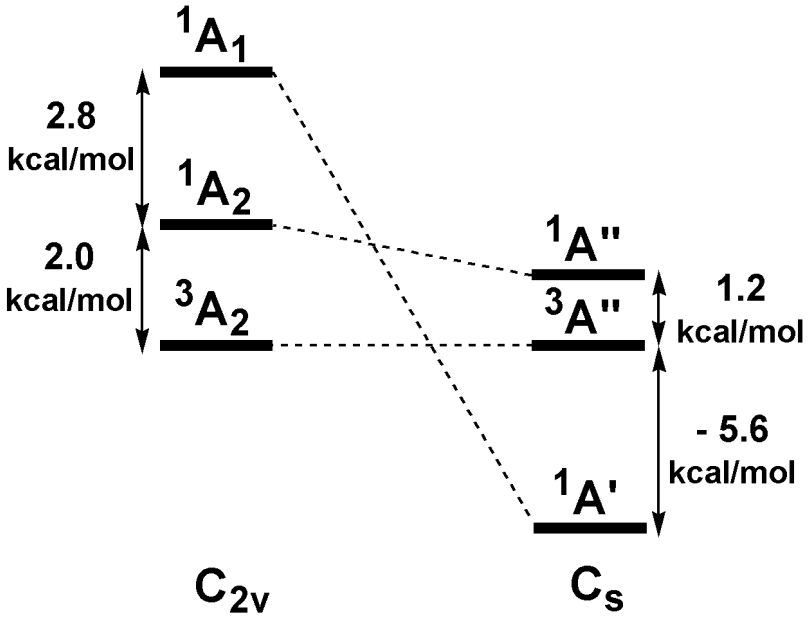


Figure 3.3. Correlation diagram for low energy singlet and triplet states of L'Cu(NPh) in C_{2v} (left) and C_s (right) symmetry.

Table 3.1. CASSCF(10,10)/CEP-31G(d) Optimized C_{2v} Geometries for (β-diketiminate)Cu(NPh)[a]

State	E_{rel} (kcal/mol)	Cu-N (Å)	N-C_{ipso} (Å)	Cu-N-C_{ipso} (Å)
$^3A_2$	0.0	1.866	1.368	180.0
$^1A_2$	2.0	1.880	1.364	180.0
$^1A_1$	4.8	1.810	1.341	180.0
$^3A_1$	14.2	1.890	1.308	180.0

[a] Calculations were carried out with phenyl substituent perpendicular to the plane of the β-diketiminate ligand.

3.3.2.2. C_s Symmetry

The singlet-triplet energy splitting calculated at C_s geometries is likely to be less than the foregoing estimates as the DFT evidence (see section 3.3.1) suggests that singlet copper nitrene has an appreciably more bent nitrene coordination mode than the corresponding triplet. For this reason, utilizing the 10-orbital, 10-electron CASSCF methodology, (β-diketiminate)Cu(NPh) was optimized under C_s symmetry for singlet and triplet states of both A' and A" symmetry. *The closed-shell ^{1}A' state is now predicted to be the ground state for L'Cu(NPh)!* Relative energies (kcal/mol) of the other symmetry states are: 6 (^{3}A"); 7 (^{1}A"); 12 (^{3}A'), Figure 3.3. While the A' states are bent (Table 3.2) the A" states return to linear nitrene coordination (150° was used as a starting guess as a compromise between sp and sp^2 hybridization of the nitrene nitrogen) upon geometry optimization. The ^{3}A" and ^{1}A" states correspond to the $^3A_2$ and $^1A_2$ states, respectively, when the calculations were done under C_{2v} symmetry.

Table 3.2. CASSCF(10,10)/CEP-31G(d) Optimized C_s Geometries for (β-diketiminate)Cu(NPh)[a]

State	Erel (kcal/mol)	Cu-N (Å)	N-C_{ipso} (Å)	Cu-N-C_{ipso} (Å)
^{1}A'	0.0	1.852	1.396	132.0
^{3}A"	5.6	1.843	1.335	179.7
^{1}A"	6.8	1.851	1.332	180.0
^{3}A'	11.7	1.994	1.401	124.1

[a] Calculations were carried out with phenyl substituent perpendicular to the C_s plane of the molecule as this conformation was found to be most stable in DFT geometry optimization.

As with multireference calculations on NPh,[42] differences are seen in the CASSCF-optimized bond lengths from the ipso carbon to the nitrene nitrogen, Table 3.2. Furthermore, interesting variations are seen in copper-nitrene nitrogen distance and copper-nitrene-ipso carbon angle for the different symmetry states of C_s-(β-diketiminate)Cu(NPh). The lowest energy singlet (^{1}A') and triplet (^{3}A") states of CASSCF(10,10)/CEP-31G(d) geometry optimized L'Cu(NPh) are bent and linear, respectively, a result consistent with the DFT geometry optimizations of these nitrenes. Apart from the high energy ^{3}A' state, Cu-N bonds to the nitrene are little changed from 1.84 - 1.85 Å, Table 3.2. The nitrene nitrogen to ipso carbon distance for the ^{1}A' ground state is ~0.06 Å longer than the analogous bond in the ^{3}A" state. Indeed, both A' states have longer N-C_{ipso} bond lengths than the A" states, regardless of spin state.

The ^{1}A' state is highly multiconfigurational in nature as evidenced by the computed natural orbital occupation numbers (NOONs) of $(\pi_{CuN})^{1.5}(\pi_{CuN}*)^{0.5}$, Figure 3.4, values far removed from the single determinant occupation numbers of 2 and 0, respectively. Hence, this state is not expected to be well described at single determinant (*e.g.*, DFT) levels of theory. Analysis of the natural orbitals from the CASSCF calculation of ^{1}A'-(β-diketiminate)Cu(NPh) supports some degree of multiple bonding character for the copper-nitrene bond, Figure 3.4.

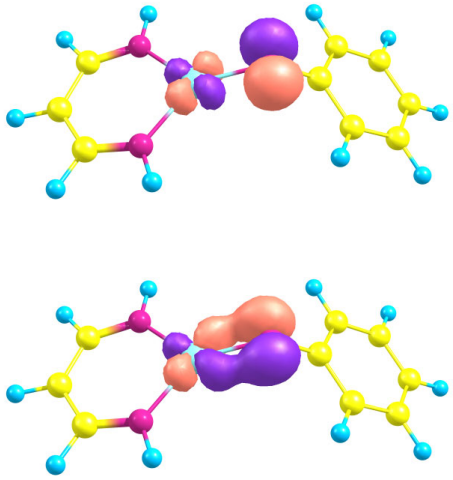


Figure 3.4. The π_{CuN} (bottom) and π_{CuN}* (top) natural orbitals for ^{1}A' state of (β-diketiminate)Cu(NPh). The natural orbital occupation number is 1.5 e- for π_{CuN} and 0.5 e- for π_{CuN}*.

The A" (~A_2) states (both singlet and triplet) of (β-diketiminate)Cu(NPh) are quite different in their character from the A' states, displaying correlation between perpendicular π symmetry orbitals that are highly polarized on either the copper or nitrene nitrogen end, Figure 3.5. In this regard, the description of the copper-nitrene bond of ^{3}A"-(β-diketiminate)Cu(NPh) within the CASCF framework is reminiscent of multiconfiguration calculations on dioxygen.[47] The nitrene nitrogen-polarized orbital is partially delocalized to the phenyl substituent as was seen in the DFT calculations on the triplet states of L'Cu(NPh), Figure 3.5 (bottom).

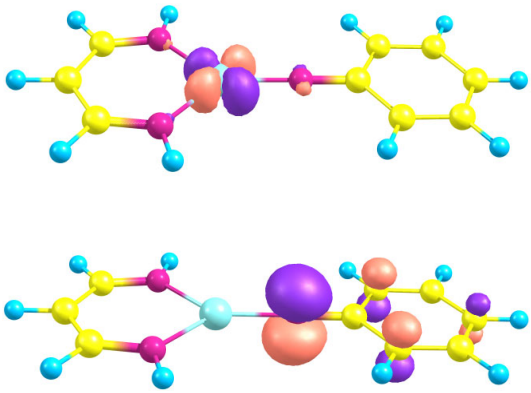


Figure 3.5. The π_N (bottom) and π_{Cu} (top) natural orbitals for ^{3}A" state of (β-diketiminate)Cu(NPh). The natural orbital occupation numbers are ~1.0 e^- for both.

3.3.3. QM/MM Calculations

As regards possible steric effects, bulky substituents were added to the β-diketiminate ligand: two mesityl and two methyl substituents ($L^{Me,Mes}$), while the nitrene substituent was set as 3,5-dimethylphenyl.[48] Figure 3.6 depicts the optimized structures (with no symmetry constraints) of singlet (^{1}A) and triplet (^{3}A) states using the hybrid CASSCF(10,10)/CEP-31G(d):MM3 method.

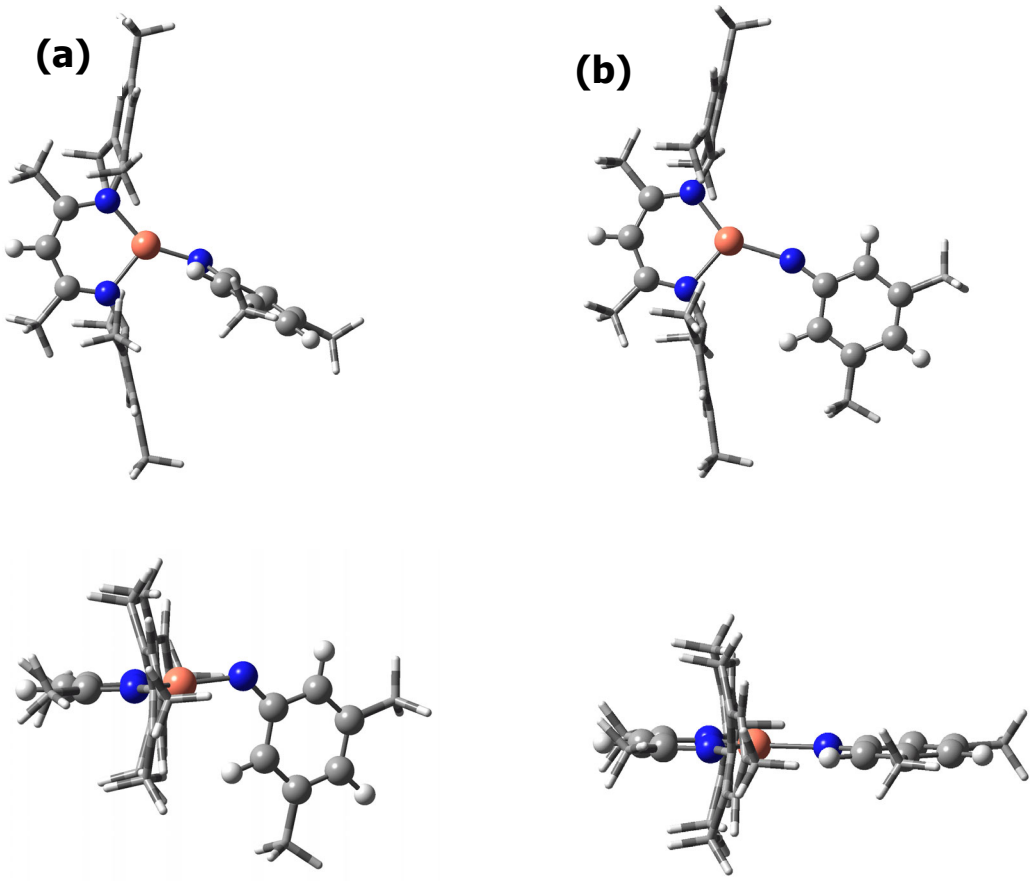


Figure 3.6. Top (top) and side (bottom) views of ($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$) optimized geometries using a QM/MM approach. The QM region (CASSCF(10,10)/CEP-31G(d)) is illustrated in ball and stick representation, while the MM region (MM3 force field) is illustrated in stick representation. (a) singlet state; (b) triplet state.

Optimized geometries of ($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$) using QM/MM hybrid methods show the same structural patterns as for smaller models: the triplet is close to linear (Cu-$N_{nitrene}$-C_{ipso} = 160.7°), while the singlet is markedly bent (Cu-$N_{nitrene}$-C_{ipso} = 130.4°). Still, the singlet-triplet energy splitting of 21.5 kcal/mol is larger than in previous QM calculations on models, caused apparently by a less favorable steric interaction between the nitrene and Cu(β-diketiminate) substituents in the triplet state.

The NOON's for the singlet ($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$) are very similar with those obtained for the ^{1}A' ground state of L'Cu(NPh). However, in the case of triplet ($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$) the natural orbitals show that both pπ orbitals of the nitrene nitrogen are occupied with ~1 electron each. Thus, the QM/MM calculations suggest that the increase of steric pressure from the β-diketiminate ligand will tend to destabilize the triplet relative to the ground state singlet of copper-nitrenes.

3.4. Summary and Conclusions

Copper-nitrene complexes are of considerable interest as intermediates in nitrene transfer although they have thus far eluded structural characterization. Thus, DFT, CASSCF and QM/MM methods in conjunction with effective core potential basis sets are employed to study the bonding and structure of such species. The nitrene complexes studied are of the form (β-diketiminate)Cu(N-aryl)[43,44,45,48] and both small and large models are considered. The "small" model (L'Cu(NPh)), where all substituents on the β-diketiminate and phenyl groups are replaced with hydrogen atoms, was investigated under both C_{2v} and C_s symmetries. The larger model (($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$)) possesses methyl and mesityl substituents on the β-diketiminate and a 3,5-dimethyl-phenyl substituent on the nitrene nitrogen. Of primary interest in this research is whether the electronic ground state of these copper-nitrenes is singlet or triplet, and the nature of the bonding character between the copper and nitrene.

While DFT calculations suggest a triplet ground state for L'Cu(NPh), more expensive CASSCF methods on L'Cu(NPh) and QM/MM calculations on ($L^{Me,Mes}$)Cu(N-3,5-

$C_6H_3Me_2$) indicated a *singlet ground state*. In all cases, singlets are found to possess a bent geometry around the nitrene nitrogen. Singlet copper-nitrenes are highly multiconfigurational as indicated by natural orbital occupation numbers of $(\pi_{CuN})^{1.5}(\pi_{CuN}^*)^{0.5}$. Hence, it is not surprising that the description of the ^{1}A' state is tenuous by single determinant methods, even a correlated technique such as density functional theory.

Analysis of wavefunctions from the CASSCF and QM/MM calculations revealed the existence of a closed-shell singlet, with some degree of multiple bonding character for the copper-nitrene bond. Indeed, in view of previous research on alkylidene/carbine,[46,49] the singlet ground state for these copper-nitrenes is itself an indicator of some degree of copper-nitrene covalent bonding for the singlet states.

The triplet states of (β-diketiminate)Cu(N-aryl), which are low energy exited states, show a completely different bonding pattern as compared to their singlet congeners. First, triplets are found to possess a linear geometry. Second, the CASSCF calculations on L'Cu(NPh) indicate that the unpaired electrons reside in highly polarized π orbitals on Cu and $N_{nitrene}$ atoms. DFT calculations on L'Cu(NPh) and QM/MM calculations on ($L^{Me,Mes}$)Cu(N-3,5-$C_6H_3Me_2$) indicate that the majority of spin density resides on the nitrene nitrogen with some delocalization of spin density to the aryl substituent. Regardless, the triplet states of copper-nitrenes display minimal evidence of copper-nitrene covalent bonding.

In light of their electronic and molecular structural differences, it is interesting that bent/singlet and linear/triplet copper-nitrenes possess CuN bond lengths that differ by a few picometers. The hypothesis is that this results from a cancellation of effects, *i.e.*, the shorter CuN distance expected from the sp-hybridized nitrogen in linear/triplet L'Cu(NPh) versus sp^2-hybridized bent/singlet L'Cu(NPh) countermands the CuN bond shortening of the latter due to partial multiple bonding character.

The shortened $N\text{-}C_{ipso}$ bond lengths and delocalization of the $N_{nitrene}$ π orbitals onto the phenyl substituent for both the singlet ground state and low energy triplet excited state suggest that $Cu\text{-}N_{nitrene}$ bonding, and hence their resulting reactivity, may be substantially modified by using various substituted phenyl groups. Similar observations with regard to NAr π-delocalization were made in a general study of the more prevalent early-middle nitrene/imido complexes.[50]

CHAPTER 4

MULTICONFIGURATIONAL THEORY ON A BINUCLEAR Cr(II) SILOX[51,ψ]

4.1. Introduction

Wolczanski and coworkers synthesized and characterized a chromium (II) dimeric species $[Cr(\mu\text{–}OSi^tBu_3)(OSi^tBu_3)]_2$, which involves a three-center geometry for the two bridging oxygen atoms and the two metallic centers.[51] The crystal structure (Figure 4.1) revealed two inequivalent molecules in the asymmetric unit cell that have very small geometric differences, *e.g.*, 2.64 and 2.68 Å for the Cr…Cr distances. The core of the molecule (the chromium and bridging oxygen atoms) was found to be considerably puckered (CrOCrO dihedral angle ~ 38°) with distorted trigonal geometries on the central atoms due to the short Cr…Cr distance. This short distance suggests that there is interaction between the two chromium centers. Therefore, this work investigates the extent and nature of any Cr…Cr interaction in the Wolczanski chromous dimers.

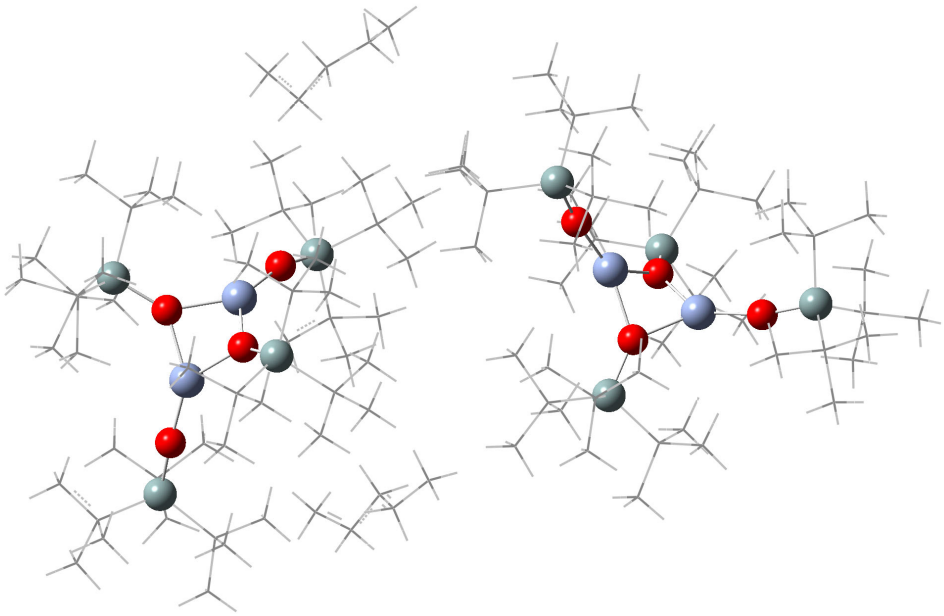


Figure 4.1. Molecular structure of $[Cr(\mu\text{–}OSi^tBu_3)(OSi^tBu_3)]_2$.

A three-coordinate chromium(II) dimer with two bridging di–tert-butylmethoxide ligands and two terminal tri–tert-butylmethoxide ligands[52] was reported by Murray *et al.*, but the Cr…Cr internuclear distance was much longer (3.075 Å) than in the Wolczanski complex and the central core of this earlier complex was planar. The magnetic moment of the latter complex was not discussed by Murray *et al.* Two similar three-coordinate chromium(II)-amide dimers, $[Cr(\mu\text{–}N\text{-}i\text{-}Pr_2)(N\text{-}i\text{-}Pr_2)]_2$[53] and $[Cr(\mu\text{–}NRR')(NRR')]_2$[54] (R = 3,5-$Me_2C_6H_3$, R' = adamantyl), were reported by Edema *et al.* and Ruppa *et al.*, respectively. Both complexes have puckered Cr_2N_2 cores and Cr…Cr distances of 2.87 Å and 2.85 Å, respectively. However, the magnetic moments showed lower values (μ_{eff} = 2.30, and 2.47 μ_B respectively) than expected (μ_{eff} = 2.83

μ_B for two unpaired electrons per dimer), suggesting an antiferromagnetic interaction between the two metallic centers.

4.2. Computational Methods

All *ab initio* quantum calculations were carried out using the GAMESS package.[37] The calculations were applied to the $[Cr(\mu\text{–}OH)(OH)]_2$ system (Figure 4.2) using the crystallographic coordinates of the full model $[Cr(\mu\text{–}OSi^tBu_3)(O\text{-}Si^tBu_3)]_2$. The geometry of the model was modified slightly to yield C_{2v} symmetry to aid in the electronic structure analysis. Both MCSCF[13] and MRMP2 (multireference Møller-Plesset second order perturbation theory)[55] methods with two selections of basis sets were employed. First, the CEP-31G[41] effective core potential basis set on all atoms with polarization functions on oxygen atoms (Basis Set I) was employed. Second, the CEP-31G basis set with polarization f[56] and (n+1)p[57] functions on chromium and the 6-311G (d,p)[58] basis set on oxygen and hydrogen atoms (Basis Set II) was employed to assess basis set effects.

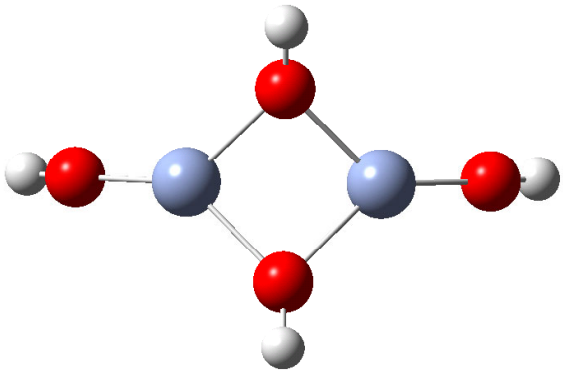


Figure 4.2. Structural geometry of the $[Cr(\mu\text{–}OH)(OH)]_2$ model with C_{2v} symmetry.

4.3. Results and Discussion

Preliminary single point energy calculations were applied at the CASSCF/Basis Set I level of theory for all plausible spin multiplicities (1, 3, 5, 7, and 9) and active spaces symmetries (A_1, A_2, B_1, and B_2). The active space included the eight lowest energy MOs derived from the Cr d orbitals and the eight d electrons from the two chromous ions (*i.e.*, a single-shell active space). The predicted low energy states were $^1A_1$, $^3B_1$, $^5A_1$, $^7B_1$ and $^9A_1$, Table 4.1, with the next lowest state being 57 kcal/mol higher. Hence, only these five low lying electronic states were studied in the subsequent calculations: MCSCF/Basis Set II and MRMP2/Basis Set I(II) for the single shell active space, and MCSCF/Basis Set I for the double shell active space. The double shell active space incorporated the single shell active space and the next eight lowest energy MOs subtending the same irreducible representations as the eight MOs in the single shell active space.

Table 4.1. Relative energies (kcal/mol) of the $[Cr(\mu OH)(OH)]_2$ model for the different spin states and the calculated magnetic moment.

Electronic state	MCSCF (single shell)		MCSCF (double shell)	MRMP2 (single shell)	
	Basis set I	Basis set II	Basis set I	Basis set I	Basis set II
$^1A_1$	0.00	0.00	0.00	0.00	0.00
$^3B_1$	0.29	0.29	0.35	0.44	0.44
$^5A_1$	0.87	0.85	1.07	1.32	1.31
$^7B_1$	1.73	1.68	2.13	2.68	2.63
$^9A_1$	2.84	2.76	3.56	4.58	4.48
μ_{eff}	3.1 μ_B	3.2 μ_B	2.8 μ_B	2.4 μ_B	2.4 μ_B

The calculated results in Table 4.1 showed that the method (MCSCF *vs.* MRMP2) is more important than the basis set selection (Basis Set I *vs.* Basis Set II) with respect to the state energy splitting. However, all methods indicate that the $^1A_1$ is the most stable state (Table 4.1). Moreover, both MCSCF and MRMP2 show the same trend for the relative energies, *i.e.*, energy increases slightly with spin multiplicity. Assuming a Boltzmann distribution for the various low energy multiplets, the calculations at the various levels of theory indicate that the ground state electronic structure is on average 58% $^1A_1$, 31% $^3B_1$, and 9% $^5A_1$ with smaller amounts of the $^7B_1$ and $^9A_1$ states.

Including the spin degeneracy, the calculated Boltzmann distribution reveals that the major contribution to the magnetic moment is from the triplet and quintet states in a roughly 2:1 ratio: 28% for $^1A_1$, 44% for $^3B_1$, 22% for $^5A_1$, 5% for $^7B_1$, and 1% for $^9A_1$.

Table 4.2. Active space MOs electronic occupation calculated with MCSCF/Basis set I.

Electronic state	Number of doubly occupied MOs				
	4	3	2	1	0
$^1A_1$	27%	9%	53%	1%	4%
$^3B_1$	-	52%	14%	23%	-
$^5A_1$	-	-	85%	4%	9%
$^7B_1$	-	-	-	100%	-
$^9A_1$	-	-	-	-	100%

The different states of the chromous dimer are highly multireference. The natural orbital population shows different occupations for the active space MOs (Table 4.2). This variety of electron configurations is possible because the d orbitals of Cr atoms are

densely packed in the Wolczanski dimer, ~ 50.7 kcal/mol for the $^9A_1$ state. Interestingly, the orbital occupancy for the $^1A_1$, $^3B_1$, and $^5A_1$ states suggests a significant antiferromagnetic arrangement of the chromium d electrons. An illustration of the antiferromagnetic coupling for the $^1A_1$ state with 2 doubly occupied MOs is shown in Figure 4.3.

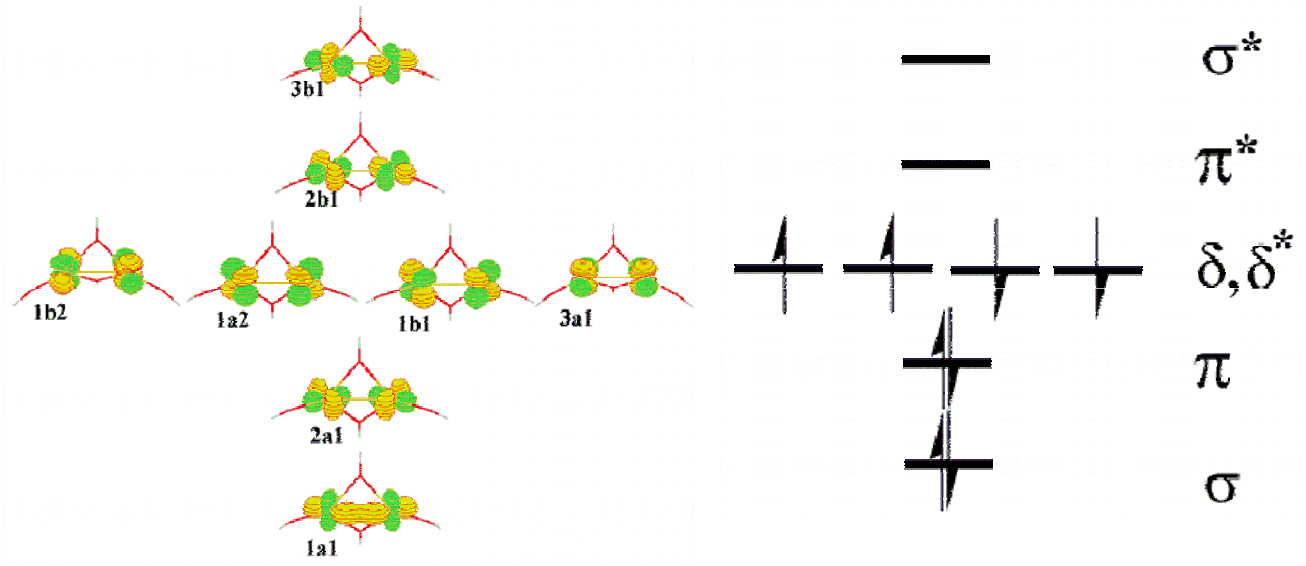


Figure 4.3. Occupation of the active space MOs and the antiferromagnetic coupling ($^1A_1$ state).

The predicted magnetic moment (Table 4.1) is method dependent, but the differences are small and all calculated values (μ_{eff} = 2.4 – 3.2 μ_B, Table 4.1) show a good agreement with the experimental value (μ_{eff} = 2.8 μ_B).

4.4. Conclusions

Single point energy calculations were performed on a symmetric C_{2v} chromium dimer using multireference electronic structure methods and effective core potentials. The geometry was derived from the crystal structure of the $[Cr(\mu\text{–}OSi^tBu_3)(OSi^tBu_3)]_2$ complex. Although monomeric chromium silox complexes have a preference for high spin,[59] the dimer complex shows a greater stability in the low spin states, suggesting some degree of Cr···Cr interaction. Five low energy electronic states were found within a range of less than 5 kcal/mol: $^1A_1$, $^3B_1$, $^5A_1$, $^7B_1$ and $^9A_1$. The chromium d orbitals remain closely packed in each state and the most significant contributions to the magnetic moment come from the triplet and quintet states. The $^1A_1$, $^3B_1$, and $^5A_1$ states involve electron configurations that arise from the antiferromagnetic coupling of the chromium d electrons.

CHAPTER 5

PREFERRED GEOMETRY OF A THREE-COORDINATE NICKEL(I) CARBONYL COMPLEX[60,§]

5.1. Introduction

Three-coordinate complexes of unsaturated transition metals display substantial reactivity and have unusual electronic structure.[61] There are few examples of three-coordinate nickel complexes.[62] Recently, Holland and coworkers isolated a three-coordinate β-diketiminate-nickel(I) complex with a carbonyl ligand that displays a T geometry, contrasting with that of isoelectronic d^9 copper (II) complexes, which have Y geometries.[60] A bulky β-diketiminate ligand (L) with two methyl and two 2,6-di-isopropylphenyl groups was used in the experimental studies. Three-coordinate nickel (I) complexes are also important from a biological point of view. The acetyl-coenzyme A synthase has two nickel centers in its active site, one of them (proximal Ni) having a low coordinate environment.[5]

5.2. Computational Methods

All calculations were carried out utilizing the Gaussian'03 package.[36] The hybrid functional B3LYP[16] was employed together with the CEP-31G[41] effective core potential basis set; d polarization functions augmented the valence basis set of main group heavy atoms. Certain Ni and Cu complexes were studied: [(L')Ni(CO)] and [(L')Cu(X)] (L' =

β–diketiminate model; X = Cl). The calculations were applied to a simplified model (the methyl and 2,6-di-isopropylphenyl groups were replaced with hydrogen atoms) as in previous simulations[62,63] of the β-diketiminate ligand.

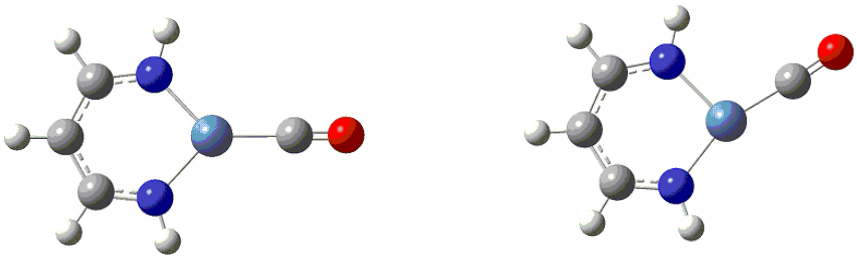


Figure 5.1. Structural models of the (L')Ni(CO) complex.

Two initial doublet ground state geometries (Figure 5.1) for each complex were fully optimized at the B3LYP/CEP-31G(d) level of theory. The first geometry had C_{2v} symmetry (Y-shape) and second geometry had C_1 or C_s symmetry (T-shape geometry). Since the preliminary unrestricted calculations showed evidence of moderate spin contamination, all the calculations were performed using the restricted open-shell Kohn-Sham formalism. The energy Hessian was calculated to confirm that stationary points are minima (*i.e.*, no imaginary frequencies) and Natural Bond Orbital (NBO[64]) analysis was applied in order to estimate the electronic distribution.

5.3. Results and Discussion

5.3.1. (L')Ni(CO) Complex

Two low-energy electronic states are found for the Y-shape geometry: $^2B_1$ and $^2B_2$, the latter being more stable by 12.8 kcal/mol. However, the calculations show that the T-shape ground state geometry is energetically favored by 7.5 kcal/mol over the lowest energy Y-shape ($^2B_2$ electronic state) geometry, consistent with the crystal structure information and the formal d^9 ground state of Ni(I).[65,66] The calculated metric data for (L')Ni(CO) are presented in Table 5.1, which shows very good agreement between the calculated T-shaped minimum and experimental results.

Table 5.1. Structural data for the calculated (L')Ni(CO) geometries.[a]

	Calc. C_{2v}		Calc. C_1	Expt.
	$^2B_1$	$^2B_2$	^2A	
Bond lengths (Å)				
Ni - C	1.873	1.867	1.798	1.771
C = O	1.165	1.163	1.166	1.141
N_1 - Ni	1.937	1.896	1.902	1.869
N_2 - Ni	1.937	1.896	1.930	1.917
Bond angles (º)				
N_1- Ni - N_2	95.25	94.48	96.62	96.42
N_1- Ni - C	132.37	132.76	159.92	158.84
N_2 - Ni - C	132.37	132.76	103.46	104.68
Ni - C = O	180.00	180.00	176.37	177.51
Improper torsion	180.00	180.00	179.88	178.48

[a] $C_{2v} \rightarrow$ Y-shape; $C_1 \rightarrow$ T-shape.

Analysis of the frontier molecular orbital energy diagrams (Figure 5.2) for the Y-shape and T-shape geometries of (L')Ni(CO) show no evidence of a Jahn-Teller effect, and that the major effect in driving the Y $\rightarrow$ T distortion is stabilization of a_2 and a_1 orbitals to form a'' and a' orbitals (assuming approximate C_s symmetry).[65,67] Investigation of NBO calculations revealed that the atomic charge distribution between the ligands and Ni center is –0.6 for L', +0.7 for Ni, and –0.1 for CO. The same result was obtained for both geometries, C_{2v}(Y) and C_1(T). The electronic distribution on the Ni atom showed 8.8 electrons in the d orbitals, very close to the ideal d^9 configuration.

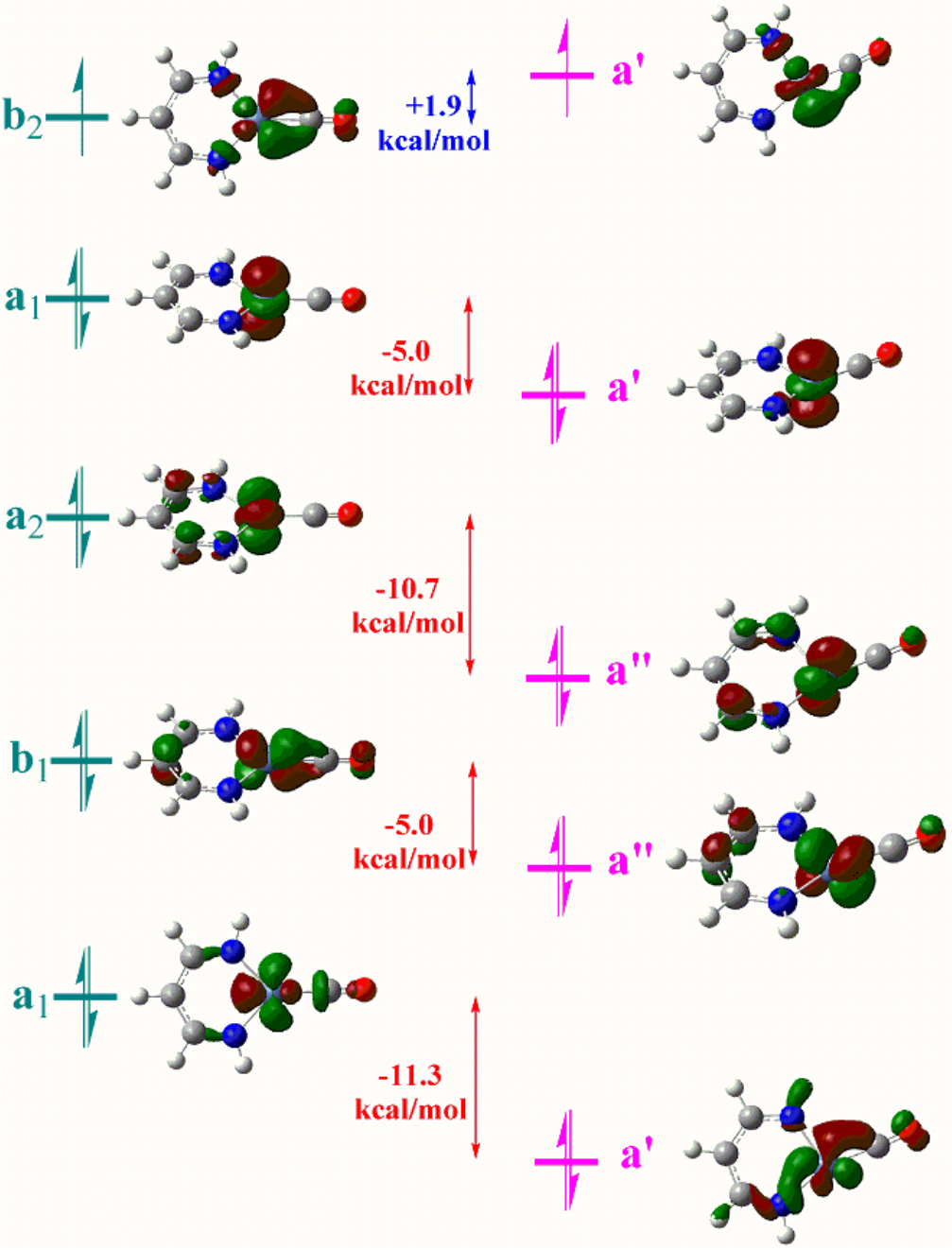


Figure 5.2. Orbital diagram showing the energy stabilization between Y and T shape.

5.3.2. (L')CuCl Complex

For the (L')CuCl complex, the results showed that the Y-shape ($^2B_2$ electronic state) and T-shape ($^2A'$ electronic state) are energetically very close, with a separation of only 1.6 kcal/mol in the favor of the C_s (T-shape) geometry. The calculation of the C_{2v} geometry generated an imaginary frequency with the chlorine atom oscillating in the plane of the molecule, a result that supports a preference for the distorted Y-shape. In this case the distortion from the Y-shape is much less pronounced than for the (L')Ni(CO) complex: the N-Cu-Cl bond angles differ by only 13.6º compared to N-Ni-CO bond angles that differ by 56.5º.

As expected, the NBO analysis shows the same charge distribution for both C_s and C_{2v} geometries of (L')CuCl: -0.5 for L, +1.2 for Cu, and –0.7 for Cl. The number of electrons in the Cu d orbitals is 9.3, which is considerably closer to a d^{10} configuration than the 8.8 d electrons calculated for the Ni complex, even though the transition metal in both complexes should formally be d^9.

Several β-diketiminate copper(II) complexes were reported[68] to have more or less distorted Y-shapes. The bond angles around the cooper center are summarized in Table 5.2. The complexes containing alkoxide (OR^-) ligands are more likely to distort to the T-shape compared to those with thiolate (SR^-) and chloride (Cl^-) ligands. The more electronegative ligands (OR^-) would withdraw increased electron density from copper, which is very likely to reduce the d electron count to the standard d^9 configuration, thus leading to a T-shape geometry for (L')CuOR.

Table 5.2. Distorted Y-shape geometries for the β-diketiminate copper complexes.

Bond angles (°)	N_1 - Cu - N_2	N_1 - Cu - L	N_2 - Cu - L
(L')CuCl	97.22	132.36	130.41
(Cl-L')CuCl	96.10	132.00	131.84
(L')Cu($SC_6H_3Me_2$)	96.76	131.56	130.80
(L')Cu($SCPh_3$)	96.77	131.19	128.23
(L')Cu($SCPh_2CH_2OMe$)	96.24	135.89	126.58
(L')Cu(OC_6H_4OMe)	96.66	140.86	122.13
(L')Cu($OC_6H_3Me_2$)	96.31	138.75	123.98
(Cl-L')Cu($OC_6H_4{}^tBu$)	96.26	145.72	117.77

5.4. Conclusion

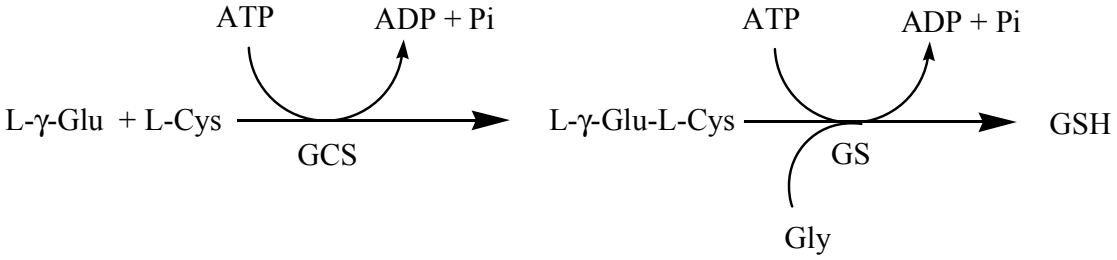
The geometry differences between the Ni and Cu β-diketiminate complexes are caused by the electron density variation in their d orbitals. Although typical electron counting rules indicate that both metals are formally d^9, NBO calculations indicate that Cu has a tendency for a d^{10} configuration, while Ni in analogous three-coordinate complexes shows a preference that is closer to d^9. This phenomenon is presumably caused by the β-diketiminate ligand, which donates more or less electrons to the metal center in response to the varying electronic demands of the other ligand (π-acid CO in the case of (L')Ni(CO) and a univalent ligand in the case of (L')CuX). The experimental data suggest that the more electronegative the X ligand (*e.g.*, OR), the closer to d^9 is the electronic configuration of the Cu ion and thus, the resulting coordination geometry is more distorted to the T-shape geometry.

CHAPTER 6

LOOP MOTION IN HUMAN GLUTATHIONE SYNTHETASE[ר]

6.1. Introduction

Glutathione (GSH) is an essential tripeptide antioxidant for many biological processes and is present in almost all bacteria, plant, and animal cells.[69] It protects cells against reactive oxygen species, is involved in amino acid transport, in metabolism of therapeutic drugs, mutagens and carcinogens, and in the maintenance of protein thiol groups and ascorbic acid in their reduced states.[70] The cellular levels of GSH are controlled by use and biosynthesis. Decreased levels of GSH have been observed in Parkinson's disease, Alzheimer's disease,[71] HIV, hepatitis C, type II diabetes, idiopathic pulmonary fibrosis, adult respiratory distress syndrome (ARDS), and cataracts.[70]



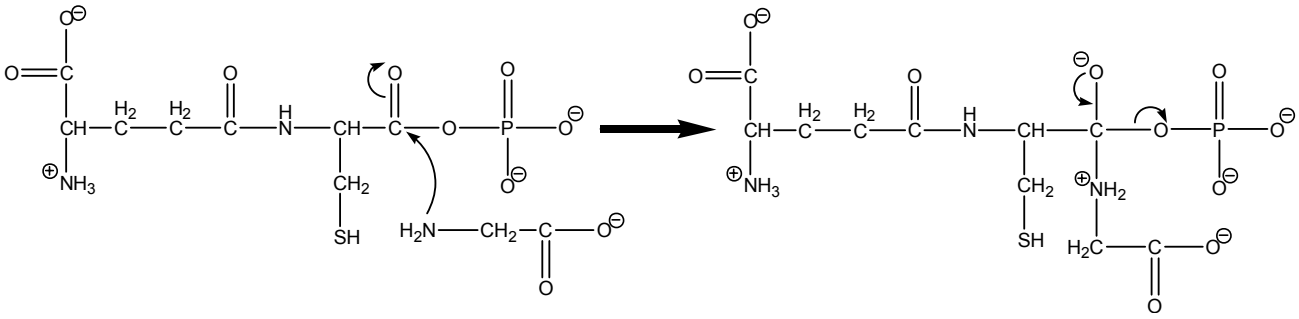
Scheme 6.1

Two ATP-dependent (Scheme 6.1) enzymatic steps[70] are necessary for GSH biosynthesis. In the first step, γ–glutamylcysteine synthetase catalyzes the formation of

---

[ר] Reprinted from Biochemical and Biophysical Research Communications 353, Adriana Dinescu, Mary E. Anderson, Thomas R. Cundari, "Catalytic Loop Motion in Human Glutathione Synthetase: A Molecular Modeling Approach." Pages 450-456, Copyright 2007, with permission from Elsevier.

L-γ–glutamylcysteine from L-glutamate and L-cysteine. The second step, which ligates γ–glutamylcysteine with glycine, is catalyzed by glutathione synthetase (GS). Both reactions most likely involve the formation of an enzyme-bound acyl-phosphate intermediate prior to the peptide bond formation.[72] Thus, in the GS-catalyzed reaction, γ–glutamylcysteinyl-phosphate is attacked by glycine (Scheme 6.2), forming an enzyme-product complex, which finally dissociates with release of GSH, ADP and phosphate (P_i).



Scheme 6.2

Several superfamilies of ATP-binding proteins are known.[73] The GS protein belongs to the Glutathione synthetase ATP-binding domain-like (ATP-grasp) superfamily, which consists of enzymes with ATP-dependent carboxylate-amine ligase activity. Superfamily members display a unique nucleotide-binding fold, referred to as an ATP-grasp fold. Proteins belonging to the ATP-grasp fold have tertiary (*i.e.*, three-dimensional) structural similarities, but have a low degree of sequence (i.e., amino acid) similarity. Although these proteins bind a variety of substrates, all ATP-grasp enzymes require ATP and are thought to involve acyl-phosphate intermediates.[74] Both eukaryotic and prokaryotic GS enzymes have an ATP-grasp fold but have different numbers of subunits as well as dissimilar sequences. In fact, human GS (hGS) is a circularly permuted[75] version of the prokaryotic enzyme that only conserves the secondary

structure elements (sequence identity is lower than 10%). Mammalian GS is a homodimer, while bacterium GS enzyme is a tetramer.

Recent research[76] has been directed toward the investigation of the conserved residues in the active site of hGS enzyme relative to the members of the ATP-grasp superfamily. Using a variety of bioinformatics methods, four highly conserved residues were found in the active site of hGS: Glu144, Asn146, Lys305 and Lys364. The function of these conserved residues was studied experimentally and computationally at the molecular level by site-directed mutagenesis. The findings revealed that the four conserved residues are necessary for optimal binding of the Mg^{2+}/ATP substrate.

Structurally conserved regions are important for enzyme function. Certain flexible loops of GS enzymes are thought to be involved in catalysis. Two loops in glutathione synthetase from *E. coli* were shown to be flexible and important for catalytic activity: the small glycine-rich loop (Gly164 to Gly167) and the large loop (Ile226 to Arg241).[77-78] The structural data of hGS indicates three loops: substrate-binding loop or S-loop (Phe266 to Ser276), glycine-rich loop or G-loop (Gln366 to Leu374) and alanine-rich loop or A-loop (Ile454 to Ala466).[79] Since the loops of hGS cover the binding site, they could be responsible for opening and closing the active site of the enzyme. For the yeast enzyme,[80] structural changes were noted between the unbound and reactant conformations, in the region of the lid domain. In this enzyme, the glycine-rich loop (analogous to the human G-loop), which is part of the lid domain, was found to become an ordered element upon ligand binding. The crystal structure of GS has been reported for only three species: *E. coli* (unbound enzyme and product complex),[81-82] *S. cerevisiae*

(unbound enzyme and reactant complex),[80] and human (product complex only).[79] This work is the first to characterize conformational changes of S-, G- and A-loops in the human GS enzyme by modeling the structurally characterized product form, as well as the as yet structurally uncharacterized reactant and unbound forms of the hGS enzyme.

The computational methods used in this research employ gas phase calculations, based on the assumption that the active site of the enzyme is isolated from the solvent. Even though the enzyme conformations studied here are merely theoretical, they represent a powerful approach for predicting and suggesting future experiments.

6.2. Computational Methods

As in a previous theory-experiment study of hGS and its mutants,[76] the available X-ray structure for hGS (product form)[83] was used to initiate the molecular dynamics (MD) search for lowest energy conformations. All water molecules were removed, hydrogen atoms were added, and the system was energy minimized. The unbound and reactant enzyme conformations of hGS were modeled by (i) removing the ligands, and (ii) replacing them with the corresponding substrates (ATP, $2Mg^{2+}$, γ–glutamylcysteine, and glycine), respectively. The different states of hGS during the reaction were modeled with the MOE 2003.02 software (Molecular Operating Environment)[84] using the AMBER94 force field[18] and Marsilli-Gasteiger "Partial Equalization of Orbital Electronegativities" (PEOE) atomic charges.[85] A similar molecular dynamics protocol employed in a previous computational study of wild-type and mutant hGS was also used here;[76] it is summarized below.

For relieving steric strain in the protein, a three-phase energy minimization algorithm was employed: first the steepest descent algorithm (RMS gradient < 1000), then the conjugate gradient technique (RMS gradient < 100), and finally the truncated Newton method (RMS gradient < 0.01). Molecular dynamics simulations with constant NVT and a time step of 0.001 ps were employed in order to find the lowest energy conformation of each enzyme state. Initially, the temperature was raised progressively from 0 to 300 K in 1 ps. The system was then equilibrated at 300 K for 8 ns during which the atomic coordinates were saved every 1000 ps. Finally, the saved geometries were energy minimized with the AMBER94 force field[18] and the lowest energy conformation among these was stored for further analysis. For consistency, the enzyme-product complex conformation was also modeled using an MD simulation of 8 ns in length. Simulation times of 8 ns have been chosen based on previous tests of simulation convergence and equilibration.[86]

Structural investigations were carried out for all three modeled states of human glutathione synthetase enzymes. The root mean squared (RMS) deviations between different conformations of the enzyme were calculated using Swiss Pdb-Viewer.[87] For comparative analysis, several eukaryotic GS sequences were obtained from the Swiss-Prot and TrEMBL Database,[88] since only a few GS crystal structures have been reported in the Protein Data Bank.[83] The multiple sequence alignment was performed using the ClustalW[89] method, while the pairwise structural alignment (when it was applicable) was performed using the K2[90] method. Hydrogen bonds were estimated using the following

criteria: the distance between the hydrogen and acceptor atoms is shorter than 3.5 Å and the angle formed by the donor, hydrogen, and acceptor atoms is greater than 90°.

6.3. Results and Discussion

6.3.1. Sequence Alignment of Human Glutathione Synthetase

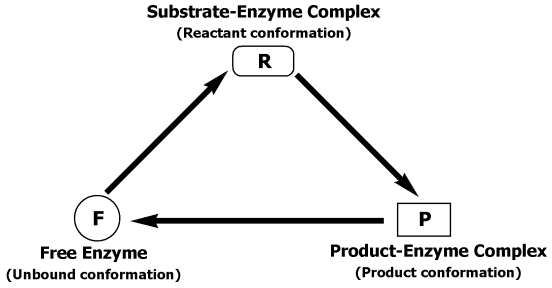
Based on structural and sequence conservation, the study of loops can provide important information about the catalytic role of the loop residues. Multiple sequence alignment analysis (Figure 6.1) for several eukaryotic species of GS shows that S-loop and A-loops residues are conserved only for mammals. The A- and S-loops residues also have lower identities (46-100% and 36-82%, respectively) within eukaryotic species, while the G-loop residues are highly conserved (90-100% identity) for all eukaryotic species, including lower species, such as baker's yeast and fission yeast. Similar to the G-loop, the glycine-rich portion of the A-loop (Gly459 to Gly464) is highly conserved for all eukaryotic species.

		S-loop			G-loop			A-loop	
Human	266	FRDGYMPRQYS	276	366	QREGGGNNLY	375	454	IEHADGGVAAGVA	466
Rat	266	FRDGYMPSQYN	276	366	QREGGGNNFY	375	454	IEHADGGVAAGVA	466
Mouse	266	FRDGYMPSQYN	276	366	QREGGGNNLY	375	454	VEHADGGVAAGVA	466
Afr. clawed frog	266	FRTGYVPQDYT	276	366	QREGGGNNLY	375	454	IEHSDGGVAAGVA	466
Tomato	341	FRAGYAPSDYH	351	437	QREGGGNNIY	446	526	SSSDEGGVAAGFA	538
Indian mustard	325	YRSGYTPRDYP	335	421	QREGGGNNIY	430	510	SSSDEGGVAAGYA	522
Mouse-ear cress	334	FRSGYTPNDHP	344	430	QREGGGNNIY	439	519	ASSDEGGVAAGFG	531
Baker's yeast	284	YRTGYTTTDYT	294	384	QREGGGNNVY	393	471	NTSNEGGVAAGFG	483
Fission yeast	290	YRVGYALDDYP	300	389	QREGGGNNTY	398	477	KKTNEGGVATGYA	489

Figure 6.1. Sequence alignment of glutathione synthetase enzymes. Only the loop residues of human GS and their equivalents in other species are shown. Shaded areas indicate the amino acid identities. The ClustalW sequence alignment between human and baker's yeast enzymes was confirmed by K2 structural alignment.

6.3.2. Analysis of Reactant, Product and Free Forms of hGS

The motion of hGS loops during the reaction cycle (*i.e.*, transition from free enzyme to reactant to product form, Scheme 6.3) is studied by employing an MD conformational search, similar to previous research.[76] Starting from the product crystal coordinates, the product (P), reactant (R), and free (F) enzyme conformations are modeled using molecular mechanics. These three calculated enzyme geometries are then superimposed (Figure 6.2) and the conformational changes are analyzed by determining the RMS deviations (Tables 6.1, 6.2, 6.3, and 6.4) in corresponding amino acid residues.



Scheme 6.3

The calculated backbone RMS deviations for the overall enzyme structures (Table 6.1, first row) are 1.96 Å (F versus R conformation), 2.83 Å (R versus P conformation), and 2.51 Å (P versus F conformation). Thus, these results for hGS show no significant change in the global tertiary structure of the enzyme during the reaction cycle. The calculated side chain RMS deviations are 2.60 Å (F versus R conformation), 3.38 Å (R versus P conformation), and 3.09 Å (P versus F conformation). The structural alignment results (Table 6.1, first row) for the three enzyme conformations show that side chain atoms display *ca.* 0.5 Å greater RMS deviation for each form than the deviation for the

backbone atoms. This trend of larger deviations for the side chain atoms is also valid for the three G, A, and S loops as well, whether treated together as an *aggregate* (Table 6.1, second row) or *individually*, with a minor exception for the A-loop. These results thus suggest that side chain motion of the G-, A-, and S-loop residues could be important for enzyme activity.

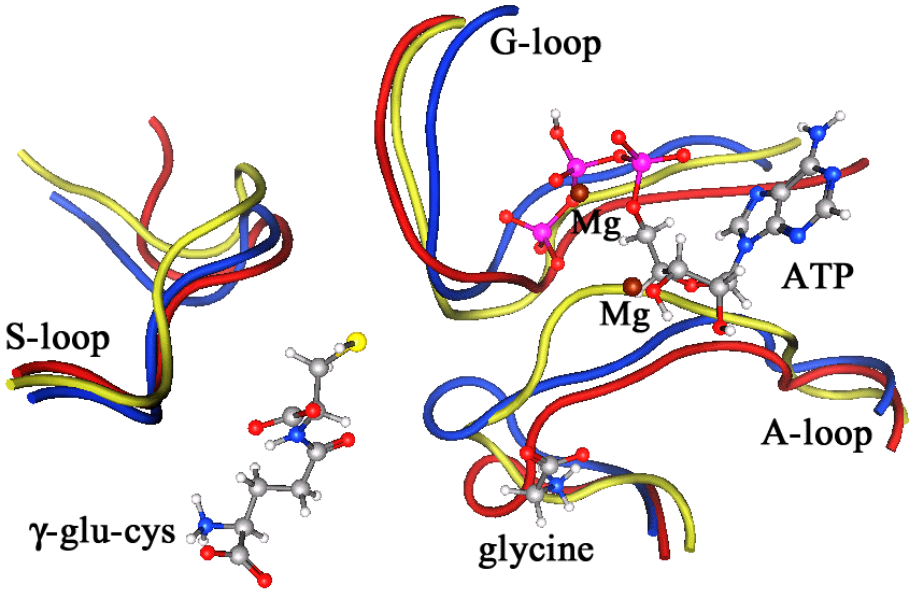


Figure 6.2. Calculated structural superposition of the G, A, and S loops during the reaction cycle. The unbound (yellow) and reactant (blue) conformations are superimposed against the product (red) enzyme complex. The substrates are included for visualization of their position relative to the loops.

Table 6.1. RMS deviations of hGS loops.[a]

Enzyme region	RMS (Å)					
	Unbound → Reactant[b]		Reactant → Product[c]		Product → Unbound[d]	
	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms
Overall enzyme	1.96	2.60	2.83	3.38	2.51	3.09
GAS-loops	2.43	3.08	2.54	3.60	2.67	3.19
G-loop	2.41	3.50	2.82	5.08	2.80	3.52
A-loop	2.75	3.59	2.41	2.40	2.75	3.33
S-loop	1.99	2.41	2.41	2.84	2.69	2.87

(a) Conformational changes were measured using the calculated RMS after performing a structural alignment between different enzyme states.
(b) Conformational changes between the calculated unbound and reactant-complex enzyme states.
(c) Conformational changes between the calculated reactant-complex and product-complex enzyme states.
(d) Conformational changes between the calculated product-complex and unbound enzyme states.

The largest RMS values for both backbone and side chain atoms of the three loops are obtained for the G- and A-loops. The smallest RMS deviations were obtained for the S-loop (Table 6.1, fifth row), which are close to the RMS values for the overall enzyme during the reaction cycle. The A-loop shows the largest RMS deviations only during the F→R transition (2.75 Å for backbone atoms and 3.59 Å for side chain atoms, Table 6.1, fourth row). Notably, the G-loop shows the largest RMS deviations for the backbone and side chain atoms during the R→P and P→F transitions (Table 6.1, third row).

In summary, the calculated RMS deviations suggest that both the G- and A-loops of hGS have a significant contribution to substrate binding, peptide bond formation and

product elimination due to their larger motion during the reaction cycle. The S-loop plays a lesser role because of its smaller RMS deviations. Since the A loop shows the largest deviations between the F and R structures, the A-loop is more involved in binding the substrates. Similarly, since G-loop residues show the largest deviations in the other two transitions (from R to P and from P to F) with a significant RMS deviation between the R and P conformations, the G-loop is more involved in the product formation and possibly in product elimination. Furthermore, it is noticeable that specific residues within the G-, A-, and S-loops have large RMS deviations (Tables 6.2, 6.3, and 6.4), suggesting that these residues change their conformation to accommodate the substrates inside the active site, participate in catalysis, or release the product. The most significant among these residues are discussed in the following sections.

6.3.3. Function of G-loop Residues in hGS

The reorganization of the G-loop residues during the transition between the unbound, reactant and product states of hGS is shown in Figure 6.3 (see also Table 6.2 for backbone and side chain motion). During the transition from F to R conformation, the guanidyl group of Arg367 shifts from one side of the loop to the other (3.55 Å absolute deviation for the ε-N atom), presumably to allow the entry of the ATP inside the binding site of enzyme and close the active site for reaction. After the reaction is completed it moves again (3.93 Å absolute deviation for the ε-N atom during the transition from the P to F enzyme), opening the active site. The closure occurs by the disruption of a 3.06 Å hydrogen bond between the backbone oxygen atom of Arg367 and the δ-amino group of Asn372 and formation of a 2.37 Å ionic hydrogen bond

between the η-amino group of Arg367 and the γ-oxygen atom of Asn372. Finally, this ionic hydrogen bond breaks when the products are formed. Taken together, these findings suggest that Arg367 is important for active site closure.

Similarly, the carboxylate group of Glu368 also displays a larger motion as compared to the overall side chain (Table 6.2). The absolute deviation of the δ-carbon increases from 2.92 Å between the F and R, to 4.54 Å between R and P, and to 4.14 Å between P and F conformations. The MD simulations indicate that the role of Glu368 is to bind one of the Mg^{2+} ions throughout the catalytic process. Both Glu368 and Mg^{2+} orient the ATP phosphate groups towards the reaction center and then assist in product elimination. In the R structure the binding distances between the two oxygen atoms of the γ-carboxylate group and the catalytic Mg^{2+}501 are 2.10 Å and 2.36 Å, while in the P structure the distances are 2.10 Å and 2.66 Å, respectively, suggesting that the side chain of Glu368 moves together with the Mg^{2+}501 and the γ-phosphate. Thus, Glu368 is important for the chemical reaction, being responsible for charge stabilization around the reaction intermediate.

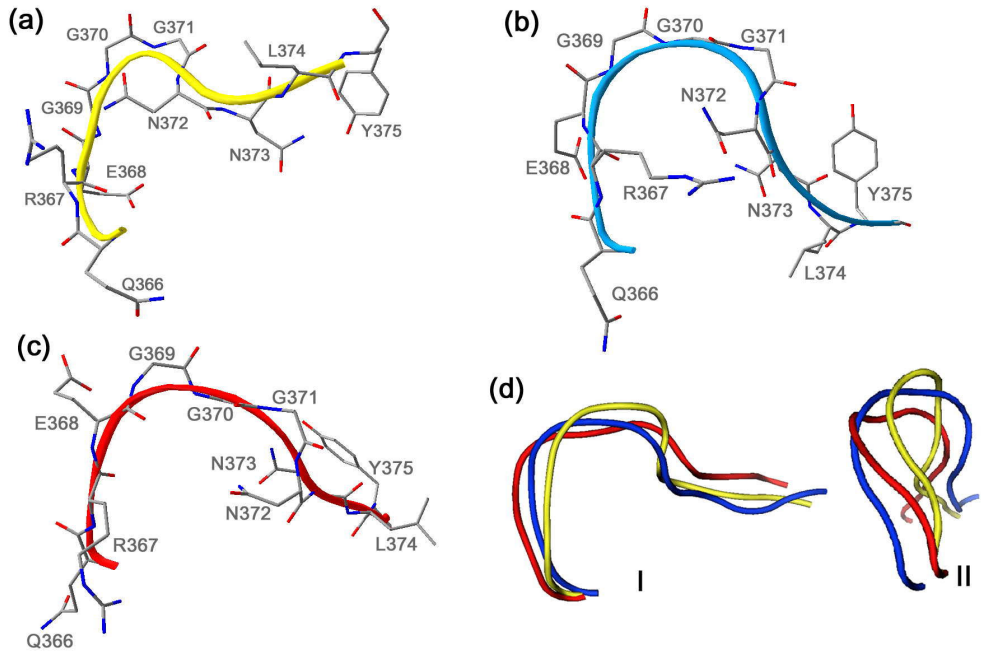


Figure 6.3. Reorganization of the G-loop residues during the catalytic cycle of hGS. The unbound (a), reactant-complex (b), and product-complex (c) enzyme structures were modeled through 8 ns of MD simulation. The G-loop residues are shown together with their backbone ribbon representation (yellow, blue and red) for each modeled conformation. Superposition (d) of the three states (I-front view and II-side view) illustrates the motion of the G-loop.

While the consecutive glycine residues (Gly369, Gly370 and Gly371) do not have a side chain, their backbone flexibility allows the substrates access to the active site. Their backbone lengths (internuclear distance between the nitrogen of Gly369 and carbonyl carbon of Gly371) elongate from 6.35 Å in the F to 7.92 Å in the R, and to 8.73 Å in the P conformation. The R structure shows that all three glycine residues form hydrogen bonds with the γ-phosphate group of ATP. The hydrogen atoms of backbone amino groups of Gly369, Gly370 and Gly371 are located in the R structure at distances of 2.37 Å, 2.74 Å and 2.94 Å respectively, from the oxygen atoms of γ-phosphate, suggesting that these residues are important for ATP binding. In the P conformation

these hydrogen bonds around the phosphate groups change so that Gly369 is linked to the γ-phosphate (2.59 Å distance), Gly370 to the β-phosphate (2.24 Å) and Gly371 to the γ-nitrogen atom of Asn373 (2.91 Å). Since Gly369 remains in contact with the γ-phosphate in both conformations (R and P), it is likely that this residue is involved in acyl-phosphate formation. Overall, Gly369, Gly370 and Gly371 residues are important for ATP binding and orientation inside the hGS active site.

The next important residues in the G-loop are Asn372 and Asn373. The Asn372 side chain is involved in hydrogen bonding with the Arg367 as described above. The γ-amino group of Asn373 forms a hydrogen bond (3.05 Å) with the ATP bridging oxygen atom between the α and β-phosphate groups in the R conformation, while in the P structure this hydrogen bond moves to the α-phosphate oxygen atom that is linked to the ribose (3.12 Å). These results suggest that both Asn372 and Asn373 reorient their side chains in the R and P conformations to facilitate the transfer of the γ-phosphate group to the substrate and possibly stabilize the transition state complex.

Other G-loop residues, which do not form hydrogen bonds with the ligands but define the ends of the loop, are Gln366, Leu374, and Tyr375. The modeled structures illustrate that the amide group of Gln366 points outside the active site and displays a small movement during the reaction. In contrast, the hydrophobic side chain of Leu374 shows large motions during the transitions from the R→P form and from the P→F form of the enzyme (RMS deviations are 7.69 Å and 7.00 Å, respectively). Its orientation in the R structure suggests that the role of Leu374 is to prevent water access inside the active site and thus, protect the acyl-phosphate intermediate, while in the P

conformation this residue opens the active site and allows the solvent inside, perhaps assisting in product release.

Table 6.2. RMS deviations of the residues located in the G loop.[a]

Residues	RMS (Å)					
	Unbound → Reactant[b]		Reactant → Product[c]		Product → Unbound[d]	
	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms
Gln366	2.94	2.25	2.74	2.79	0.77	1.90
Arg367	2.37	2.90	2.06	2.80	1.10	1.44
Glu368	2.95	3.00	1.93	4.34	2.08	3.96
Gly369	2.03	-	1.71	-	3.58	-
Gly370	2.47	-	2.51	-	4.12	-
Gly371	2.67	-	2.82	-	1.64	-
Asn372	1.98	2.62	1.89	2.67	1.86	2.66
Asn373	2.87	2.91	2.73	2.05	2.06	2.82
Leu374	1.57	2.95	2.79	7.69	2.51	7.00
Tyr375	1.71	5.33	5.31	7.79	5.09	3.38

(a) The conformational changes were estimated using the calculated RMS after performing a structural alignment between different states of the enzyme.
(b) Conformational changes between the unbound and reactant-complex enzyme structure.
(c) Conformational changes between the reactant-complex and product-complex enzyme structure.
(d) Conformational changes between product-complex and unbound enzyme structure.

The last residue in the G-loop sequence, Tyr375, displays a very large RMS deviation for the backbone atoms in the P conformation compared to the F and R

structures (5.09 Å and 5.31 Å, respectively). Additionally, the side chain RMS values show that Tyr375 moves significantly (5.33 Å between F and R and 7.79 Å between R and P). The side chain of Tyr375 is probably free to move in the F conformation but the optimized structure shows that it points towards the active site. The hydroxyl oxygen atom of Tyr375 forms a 2.73 Å hydrogen bond with the γ-amino group of Asn373 in the F conformation. When ATP binds to the enzyme, the electronic repulsion between the phenyl and adenine rings causes the movement of the phenyl ring in the opposite direction. These large motions place the phenyl group closer to the adenine ring, suggesting that Tyr375 may "propel" the ATP towards the reaction center to facilitate the formation of the acyl-phosphate intermediate. Then, during the transition from product to unbound conformation, Tyr375 side chain moves again closer to the adenine ring, probably to assist the ADP elimination.

6.3.4. Function of A-loop Residues in hGS

The reorganization of the A-loop residues during the transitions between the F, R, and P states of hGS is shown in Figure 6.4. The hydrophobic side chain of Ile454 residue moves during the F$\rightarrow$R transition (3.16 Å RMS) to shield the active site for catalysis (Table 6.3). During the transition from P$\rightarrow$F, the side chain of Ile454 moves for a second time (4.02 Å RMS), suggesting that solvent is allowed to access the active site and presumably assist in product elimination.

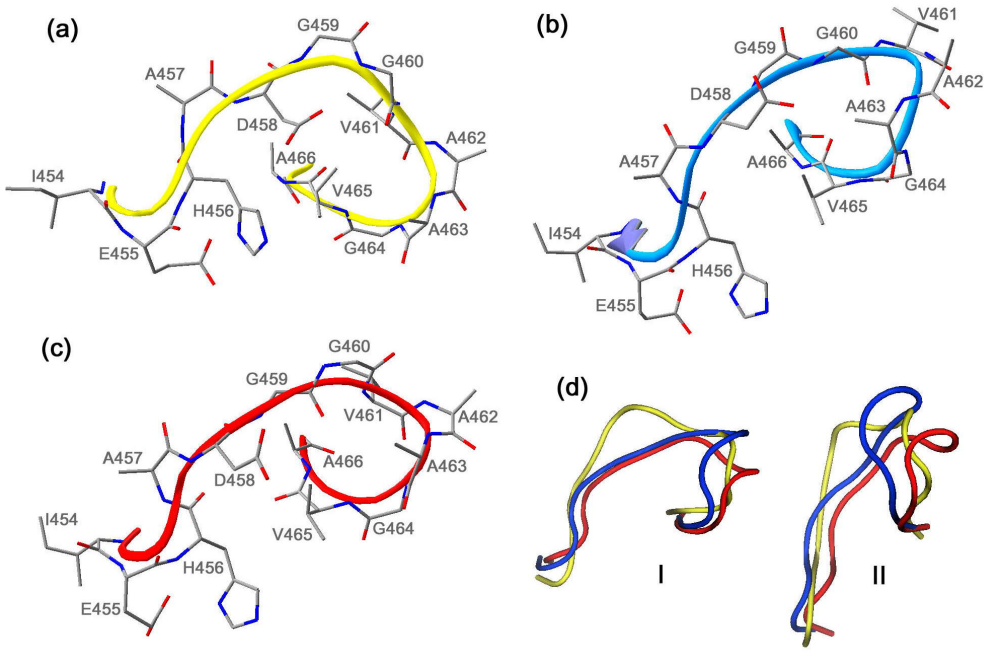


Figure 6.4. Reorganization of the A-loop residues during the catalytic cycle of hGS. The unbound (a), reactant-complex (b), and product-complex (c) enzyme structures were modeled through 8 ns of MD simulation. The A-loop residues are shown together with their backbone ribbon representation (yellow, blue and red) for each modeled conformation. Superposition (d) of the three states (I-front view and II-side view) illustrates the motion of the A-loop.

The Glu455 and His456 residues form a network of hydrogen bonds with each other that remains mostly unchanged during the catalytic process (Figure 6.5). In the F enzyme, one of the Glu455 oxygen atoms forms an ionic hydrogen bond (1.59 Å) with the imidazole ring and another ionic hydrogen bond (2.63 Å) with the backbone amino group of His456. The second oxygen atom of Glu455 forms an ionic hydrogen bond (2.44 Å) only with the imidazole ring of His456. In the R conformation these hydrogen bonds become 1.58 Å, 2.07 Å and 2.27 Å, respectively, while in the P complex they are 1.57 Å, 1.84 Å and 3.11 Å, respectively. Due to this limited conformational change during catalysis, these interactions suggest that Glu455 and His456 maintain the shape

of the A-loop throughout the catalytic process. The neighboring residue Ala457 presumably also assists in maintaining the conformation of the A-loop.

Table 6.3. RMS deviations of the residues located in the A loop.[a]

Residues	RMS (Å)					
	Unbound → Reactant[b]		Reactant → Product[c]		Product → Unbound[d]	
	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms
Ile454	1.97	3.16	1.31	1.37	2.79	4.02
Glu455	1.84	3.93	1.58	1.37	2.59	4.29
His456	1.61	3.50	1.69	1.18	1.46	3.10
Ala457	1.80	2.86	2.15	3.04	1.29	0.82
Asp458	2.16	2.06	1.29	1.49	1.81	0.87
Gly459	3.14	-	1.78	-	4.46	-
Gly460	3.49	-	2.13	-	4.53	-
Val461	4.01	6.34	4.09	5.11	3.76	5.06
Ala462	3.42	3.01	4.45	5.97	2.72	3.90
Ala463	2.95	2.12	2.52	1.96	1.24	1.68
Gly464	3.35	-	2.22	-	2.25	-
Val465	2.61	2.43	2.40	1.09	2.39	1.42
Ala466	2.06	2.57	1.04	1.54	1.45	1.86

(a) The conformational changes were estimated using the calculated RMS after performing a structural alignment between different states of the enzyme.
(b) Conformational changes between the unbound and reactant-complex enzyme structure.
(c) Conformational changes between the reactant-complex and product-complex enzyme structure.
(d) Conformational changes between product-complex and unbound enzyme structure.

The Asp458 side chain forms ionic hydrogen bonds with backbone Asp458, Gly459 and Gly460 (3.80 Å, 2.35 Å and 2.28 Å respectively) in the F enzyme, while in the other two conformations only one hydrogen bond remains between Asp458 and Gly459 (1.91 Å in the reactant complex and 1.76 Å in the product complex). The disruption of these *intra*-loop interactions is compensated for by the formation of the *inter*-loop interactions between the A- and G-loops (see interaction between G-loop and A-loop below) in the R form of hGS.

The next two residues, Gly459 and Gly460 show the highest backbone RMS deviations in the A-loop for the R→P transition (Table 6.2). Additionally, their backbones also display a significant deviation during the F→R transition. The Gly460 backbone displays a larger RMS than Gly459, suggesting that it moves to accommodate the glycine substrate. In the R complex the oxygen atom of Gly460 forms a hydrogen bond with the glycine substrate (2.05 Å), and with the backbone amide group of Ala463 (2.24 Å), while in the P structure the latter hydrogen bond is lost and Gly460 binds only the glycine moiety of GSH (2.12 Å between oxygen of Gly460 and amide group of glycine). The amide group of Gly460 also forms hydrogen bonds with the oxygen atom of the glycine substrate (3.08 Å in the R complex and 2.93 Å in the P complex). The loss of the hydrogen bond between Gly460 and Ala463 in the P complex is compensated for by the formation of other hydrogen bonds between Gly459 and Val461 (2.13 Å), Val461 and Gly464 (2.23 Å), and Ala463 and Val465 (3.02 Å). In the F enzyme the distribution of these hydrogen bonds is different: the oxygen backbone of Val461 forms hydrogen bonds with the amide backbone groups of Ala463, Gly464 and Val465

(3.34 Å, 2.47 Å, and 3.50 Å respectively). These observations show a great variety of hydrogen bonding, suggesting that the loop region between Gly459 and Val465 is very flexible and important for glycine binding to the acyl-phosphate intermediate. Additionally, this flexible region of the A-loop may perhaps play a significant role in glycine substrate recognition. The importance of the A-loop for catalysis is emphasized by sequence alignment, which identifies this flexible region from Gly459 to Val465 as highly conserved in all the eukaryotic GS enzymes (Figure 6.1).

In summary, although the conformational change of the whole A-loop is substantial, the backbone motion occurs mostly in the second part of the loop (Gly459 – Ala 466), which consists of a combination of neutral residues glycine, alanine, and valine. The presence of three glycine residues confers a larger degree of flexibility to this region. Notably, this region corresponds to the most conserved part of the A-loop and is located very close to the reaction site, all suggesting its critical role for enzyme function.

6.3.5. Function of S-loop Residues in hGS

S-loop residues are less conserved (Figure 6.1) and have less mobility than those of G- and A-loops (Table 6.4). Part of the S-loop residues is involved in binding the γ–glutamylcysteine substrate. In the R conformation, the ε-NH^+ and η-NH_2 groups of Arg267 hydrogen bond to both carboxylate oxygen atoms of the substrate (1.60 Å and 2.05 Å, respectively). Additionally, the negatively charged carboxylate group of Asp268 binds the positively charged amino group of the γ–glutamyl moiety (hydrogen bond

length is 2.07 Å). After product formation, the latter hydrogen bond is lost and only the 1.60 Å hydrogen bond between Arg267 and substrate is maintained. To compensate, Asp268 forms an intra-loop hydrogen bond with Arg273 (1.57 Å in P conformation and 2.35 Å in the F enzyme).

Table 6.4. RMS deviations of the residues located in the S loop.[a]

Residues	RMS (Å)					
	Unbound → Reactant[b]		Reactant → Product[c]		Product → Unbound[d]	
	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms	Backbone atoms	Side chain atoms
Phe266	0.83	1.31	1.23	1.44	0.64	0.93
Arg267	0.83	2.03	1.06	1.49	0.66	2.36
Asp268	1.52	3.00	0.52	2.27	1.15	1.11
Gly269	1.21	-	0.41	-	1.49	-
Tyr270	1.83	1.83	0.54	1.37	2.33	2.59
Met271	1.68	3.12	0.64	1.64	2.17	2.37
Pro272	1.85	2.03	1.89	0.89	2.71	2.33
Arg273	1.73	2.33	1.53	3.32	1.87	3.79
Gln274	2.43	2.91	3.48	5.52	1.75	3.56
Tyr275	3.35	2.93	3.88	2.65	2.68	2.67
Ser276	2.96	2.85	5.18	6.22	5.50	6.64

(a) The conformational changes were estimated using the calculated RMS after performing a structural alignment between different states of the enzyme.
(b) Conformational changes between the unbound and reactant-complex enzyme structure.
(c) Conformational changes between the reactant-complex and product-complex enzyme structure.
(d) Conformational changes between product-complex and unbound enzyme structure.

Interestingly, the most significant motion occurs at the other end of the loop (Gln274-Ser276), especially during the R→P and P→F transitions (Figure 6.2). The largest RMS deviation for the Ser276 residue (Table 6.4) is caused by the formation of a 3.51 Å hydrogen bond between its OH group and the backbone amide group of Tyr275 in the P structure. Initially, in the R conformation, Tyr275 forms a hydrogen bond with Gly269 (2.37 Å between the OH of Tyr275 and O of Gly269). The modeled structures of the enzyme and the presence of the above-mentioned hydrogen bonds suggest that the orientation of the Tyr275 side chain might influence the position of the Tyr270 phenyl ring in the active site. The latter lies on the top of the γ–glutamylcysteine substrate, revealing that Tyr270 provides a hydrophobic pocket for substrate binding. Therefore, the large movement of Ser276 appears to trigger a succession of events that align the Tyr270 residue in the active site, which protect the acyl-phosphate intermediate against hydrolysis.

6.3.6. Interaction Between G-loop and A-loop Residues in hGS

Loop residues are involved into an extensive network of ionic interactions and hydrogen bonds through their side chains or backbone atoms (Figure 6.5). As discussed above, Gly369 is involved in the γ-phosphate transfer to the substrate. In addition, the backbone oxygen of Gly369 forms hydrogen bonds to amide groups of both Val461 (2.99 Å) and Ala462 (2.44 Å) in the R complex, suggesting that this interaction covers the reaction site and thus, protect the acyl-phosphate intermediate from hydrolysis. These two inter-loop hydrogen bonds are only present in the R conformation and involve the backbone atoms of three non-charged residues. As compared to ionic

hydrogen bonds, these non-charged electrostatic interactions are weaker and therefore, more susceptible to breaking when the product is formed.

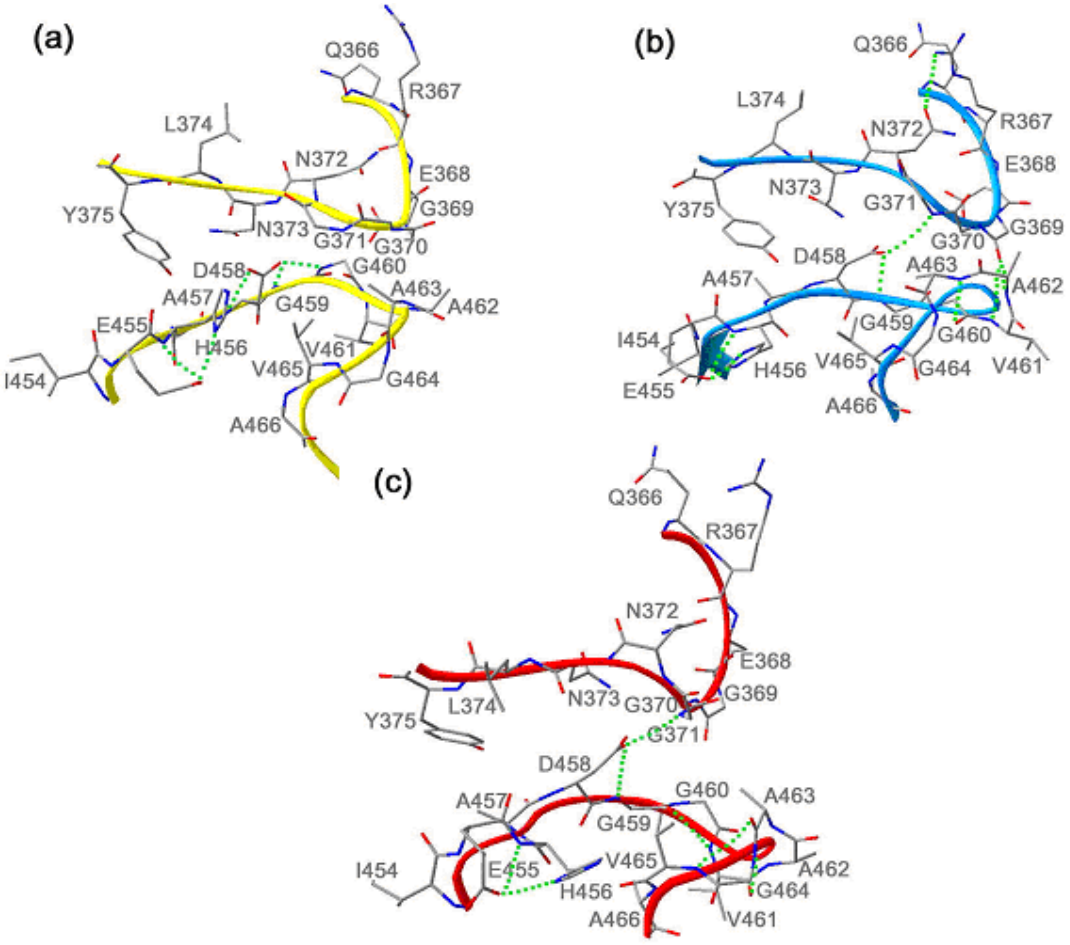


Figure 6.5. Interaction between G-loop and A-loop residues during the catalytic cycle of hGS. The unbound (a), reactant-complex (b), and product-complex (c) enzyme structures were modeled through 8 ns of MD simulation. The G and A-loop residues are shown together with their backbone ribbon representation (yellow, blue and red) for each modeled conformation. The important intra-loop and inter-loop hydrogen bonds are highlighted by green dotted lines.

Besides the non-charged electrostatic interactions between the G- and A-loops, ionic interactions are also present. The negatively charged residue Asp458 loses some *intra*-loop hydrogen bonds, which are replaced by *inter*-loop interactions in the R and P

conformations. In both R and P complexes, Asp458 forms an inter-loop ionic hydrogen bond with the amide group of Gly371 (2.83 Å and 2.21 Å, respectively). This ionic hydrogen bond is stronger and thus, important for active site closure, which is needed for catalysis.

In summary, the interaction between the G- and A-loops of hGS involves the glycine-rich portions of both loops. These regions are very close to the reaction site and show substantial flexibility. In the F conformation only the intra-loop interactions are present, while in the R and P structures some of the intra-loop hydrogen bonds are broken (especially in the A-loop, as mentioned above) and new hydrogen bonds between G- and A-loops are formed. Moreover, the glycine-rich portion (Gly369-Gly371) of the G-loop is involved in the transfer of the γ-phosphate group to the γ–glutamylcysteine substrate, while the corresponding portion (Gly459-Gly464) of the A-loop is involved in the glycine transfer to the acyl-phosphate intermediate. While many ionic hydrogen bonds are important for ligand binding, the key residues for ligand transfer and G-A loop attraction are the residues that produce less robust electrostatic interactions, but therefore have more flexibility to open and close the lid domain.

6.4. Conclusions

The crystal structure of glutathione synthetase has been reported thus far for only three species: *E. coli*, *Homo sapiens* and *S. cerevisiae*. Even though there are significant sequence and structural differences between these enzymes, the G-loop was detected in the proximity of the reaction site for all three species. The current research

shows that the amino acid sequence of the G-loop is the most strictly conserved among all eukaryotic species (lower and higher) and explains the function of G-, A- and S-loops in human glutathione synthetase.

The location of the G-loop residues very close to the γ-phosphate group of ATP and cysteinyl-glycyl moiety of GSH indicates their involvement in formation of the acyl-phosphate intermediate. Similarly, the A-loop residues are close to the glycine moiety of GSH, suggesting that part of this loop is crucial in glutathione biosynthesis. On the other side of the active site, the S-loop is responsible for binding the γ–glutamylcysteine substrate.

The side chains of Arg367 and Asn372 cover the reaction center by hydrogen bonding to each other and protect the intermediate from the solvent. The γ-carboxylate group of Glu368 coordinates the catalytic Mg^{2+} ion in the proximity of γ-phosphate in both R and P conformations. This result suggests that Glu368 participates in ATP activation and charge stabilization around the reaction intermediate. The backbone amide group of Gly369 forms hydrogen bonds with γ-phosphate of ATP in both R and P conformations. Besides the role of Gly369 in binding the substrates, its small size and remarkably flexible backbone appear to play a key role in opening a passage for glycine into the active site. The Gly460 forms hydrogen bonds with the glycine substrate in both R and P complexes, suggesting that this residue participates in catalysis. Noteworthy is the Gly459-Gly464 fragment of the A-loop, which shows great versatility of the hydrogen bonding network between the three states of the enzyme (F, R, and P

conformations). The Arg267 and Asp268 bind the γ–glutamyl part of the substrate, while Tyr270 serves as a hydrophobic lid for the GSH binding pocket.

The favorable interaction between the G- and A-loop in the R conformation suggests that these two loops provide the driving force for the active site closure, which is required for peptide bond formation. Moreover, some of the G- and A-loop residues are involved in catalysis by interacting directly with the substrates. Intriguingly, the S-loop does not interact with any of the G- or A-loop residues, but is essential for γ–glutamylcysteine binding. Thus, these residues are key players in hGS activity, and as such are promising sites for future site-directed mutagenesis experiments aimed at unlocking the mechanism of this mammalian member of the important ATP-grasp superfamily.

CHAPTER 7

GLYCINE TRIAD IN THE G-LOOP OF hGS

7.1. Introduction

Prokaryotic glutathione synthetase (GS) enzymes show little sequence identity, only 10%[79] when aligned against their mammalian counterparts. However, the glycine-rich loop motif seen in human GS (hGS) is also present in the corresponding *E. coli* enzymes. Two peptide regions from Gly164 to Gly167 and from Ile226 to Arg241 were assigned as glycine-rich loops in GS from *E. coli*, the latter (large loop) being very flexible according to the crystal structure.[82] The ATP binding site is located under the large loop structure, which proved to be essential for catalysis.[82] It has been proposed[82] that during the chemical reaction, the large loop works as a lid to protect the acyl phosphate intermediate from water and thus prevents its decomposition. Several mutations were performed on the large loop. For example, R233A and R233K mutant enzymes showed diminished activity, while R241A and R241K mutations did not significantly affect the enzyme activity.[91] Kinetics experiments suggested that the R233A and R233K mutations greatly affected the substrate binding, resulting in low affinities for γ-glu-cys and glycine. Replacement of both Pro227 and Gly240 with valine affected the catalytic role of the large loop, while the kinetic constants of the P227A, G229A, G229V, and G240A mutants were comparable to those of wild type *E. coli* GS.[77] It was suggested that mutations to valine at the hinge (P227V, G240V, and P227V/G240V double residue mutation) could restrict the loop flexibility of the *E. coli* form of glutathione synthetase.

Besides GS, there are further examples of glycine participation in loop conformational changes and ligand binding in other enzymes that have been proven to be essential for enzyme catalysis. For instance, x-ray crystallography and kinetic measurements showed that mutation of Gly38 to alanine or serine in cystolic aspartate aminotransferase[92] impaired loop structural flexibility, preventing small domain closure. Substitutions of conserved Gly76 to alanine, serine, or valine in the pepsin flexible loop[93] resulted in distorted backbone conformations that affected substrate alignment. It is well established that triosephosphate isomerase contains a flexible loop that closes the active site by interacting with the peripheral phosphate oxygen atoms of the substrate.[94] This loop was found to be important for enzyme function, which is to catalyze the interconversion between glyceraldehyde 3-phosphate and dihydroxyacetone. Truncation of four residues in the glycine region of this loop produced a mutant enzyme that perturbed the hydrogen bonds around the substrate, leading to a substantial decrease of the isomerization reaction rate. Thus, the proposed role of the glycine residues in the loop of triosephosphate isomerase is to bind and protect the enediol phosphate intermediate.

In aminolevulinate synthase (ALAS), which is involved in heme biosynthetic pathway in animals, fungi, and some bacteria, several mutations within a conserved glycine-loop indicated that two glycine residues have significant roles in cofactor (pyridoxal-5'-phosphate, PLP) binding.[95] Replacement of Gly142 by cysteine or of Gly144 by alanine, serine, or threonine resulted in considerable decreases in PLP binding affinities and catalytic efficiency of the enzyme.

Similar assessments were made for a highly mobile and conserved glycine residue located in the flexible loop of dihydrofolate reductase (DHFR),[96] which is important for the biosynthesis of purines and pyrimidines. While a G121V mutant significantly reduced ligand binding and enzyme activity, the deletion of this glycine residue also produced a conformational change as demonstrated by ligand binding kinetics that affected the formation of the Michaelis complex.

The catalytic subunit of cAMP-dependent protein kinase regulates many processes by mediating the transfer of the γ-phosphate from ATP to serine and threonine of their target proteins. Three highly conserved glycine residues in the glycine-rich loop of protein kinases are important for enzyme activity.[97,98,99] The loop acts as a nucleotide-positioning loop and prevents the access of water molecules to the active site. The function of glycine residues is directly related to ATP binding, by controlling the nucleotide affinity and phosphoryl transfer to the substrates.

Based on the structure of wild type hGS enzyme and the aforementioned literature precedents, it is likely that mutation of the three glycine residues from the G-loop (*i.e.*, G369, G370 and G371) will alter the orientations of the glycine residue backbone amides. Thus, the loss of hydrogen bonding and change in the binding of ligands (substrates and products) resulting from mutation of these glycines will lead to a loss of enzyme activity. Computational studies are thus performed to probe these hypotheses, and to guide future site-directed mutagenesis and enzyme kinetics experiments.

7.2. Computational Methods

Each of the three single point mutations, G369V, G370V, and G371V, were built separately for the reactant and product enzyme complexes obtained from the previous (see Chapter 6) simulations obtained after 8 ns of molecular dynamics. The same minimization and molecular dynamics sequence as described in Chapter 6 was used for the glycine-mutant enzymes; after minimization, 1 ns of MD simulation was employed with a sampling step of 100 ps. The lowest energy mutant conformation was selected following the energy minimization procedure applied to the ten saved geometries.

7.3. Results and Discussion

Analysis of the geometry of reactant and product complexes may provide significant details about the function of the enzyme and thus valuable leads for experimental collaborators. Two wild-type hGS conformations (enzyme-reactant and enzyme-product complexes) were modeled using 8 ns of MD simulations; the simulations were initiated from the reported product complex structure of wild-type hGS. The G369V, G370V, and G371V mutant conformations for both the enzyme-reactant and enzyme-product complexes were modeled by using an additional 1 ns of MD (see Computational Methods) to allow for reorganization of the structure in the vicinity of the single-point mutations. Comparative structure analysis for the wild-type and each of these three hGS mutants was performed with respect to ligand binding and loop-loop interaction.

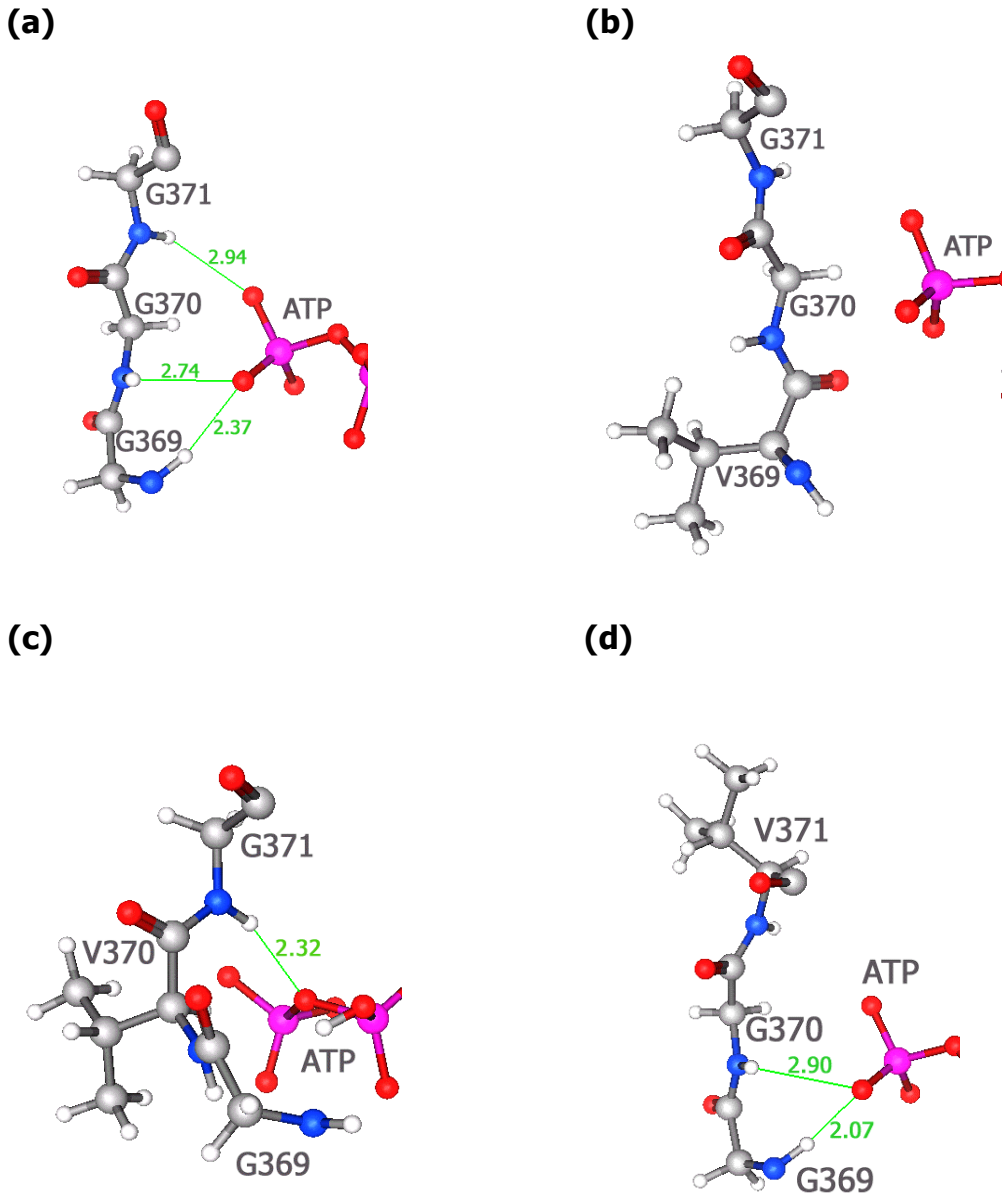


Figure 7.1. Hydrogen bonding between the γ-phosphate of ATP and G-loop residues in the reactant complex conformation. (a) wild type; (b) G369V; (c) G370V; (d) G371V.

7.3.1. Ligand Binding in the Reactant Conformation

Analysis of the wild type GS-reactant complex structure shows that all three glycine residues from the G-loop are oriented to form hydrogen bonds with the γ-phosphate of ATP (Figure 7.1.a). The amide groups of Gly369 and Gly370 form two

hydrogen bonds (2.37 Å and 2.74 Å, respectively) with the same oxygen atom of the γ-phosphate, while the amide group of G371 binds a different oxygen atom of the γ-phosphate of ATP with a hydrogen bond distance of 2.94 Å. Thus, for the native enzyme, the glycine residues are oriented so that the backbone amides form three hydrogen bonds that position the γ–phosphate of ATP (*i.e.*, the phosphate that is lost when ATP is converted to ADP and phosphate; also the phosphate that is presumably used to form the acyl-phosphate intermediate) and prepare it for catalysis.

Investigation of the modeled structure of the hGS mutant enzyme where Gly369 is replaced by valine (G369V) reveals altered orientations of amide bonds and ligands so that none of the hydrogen bonds seen in the wild type GS are able to form. While the γ-phosphate of ATP has a position similar to that of the wild type structure, the backbone amides of Val369, Gly370, and Gly371 are quite distinct in the G369V mutant. As a result, these three backbone amide groups have an unfavorable orientation to the phosphates of ATP (Figure 7.1.b) that prevents hydrogen bonding in the G369V mutant enzyme. These structural changes help to explain the experimentally observed dramatic loss of activity by the G369V mutant GS enzyme (Anderson, M. E. - personal communication).

When the glycine residue at position 370 is mutated to valine (G370V), analysis of the conformational changes shows that the orientations of ATP, of the backbones of the two glycine residues (G369 and G371), and of the mutated residue (G370V) are altered (Figure 7.1.c). In wild type hGS the three consecutive glycine residues are oriented in an extended conformation with an internuclear distance between the N of

Gly369 and the amide N of Asn372 in the reactant conformation of 7.24 Å. Analysis of the computational results for the G370V mutant GS shows that this internuclear distance (between the N of Gly369 and the amide N of Asn372) in the reactant conformation is shortened to 6.79 Å, so that the glycine region of the G-loop is slightly less extended (Figure 7.1.c). As a result, in the modeled structure, the γ-phosphate has a different orientation with an altered hydrogen bonding arrangement. Thus, instead of each of the glycine residues forming a hydrogen bond with ATP in wild type hGS (see Figure 7.1.a), in the G370V mutant enzyme only the Gly371 residue retains a hydrogen bond (2.32 Å, Figure 7.1.c) with ATP, but this hydrogen bond is now to the bridging oxygen atom between γ-phosphate and β-phosphate, rather than the γ-phosphate of ATP. Interestingly, one of the γ-phosphate oxygen atoms now forms a hydrogen bond (2.69 Å) with a glycine residue (Gly459) in the A-loop (Figure 7.3.c, see below).

Analysis of the mutant GS enzyme where glycine 371 is replaced with valine shows fewer changes in orientation and ATP binding (as compared to G369V and G370V) relative to the wild type enzyme. The orientation of its side chain prevents Val371 from forming hydrogen bonds to ATP, even though the orientation of ATP is relatively unaffected (Figure 7.1.d). The phosphates groups of ATP still have almost the same orientation in the active site of G371V as observed in the wild type enzyme. The two remaining glycine residues, 369 and 370, form hydrogen bonds to the γ-phosphate of ATP (2.07 Å and 2.90 Å, respectively) that are similar to those present in the wild type.

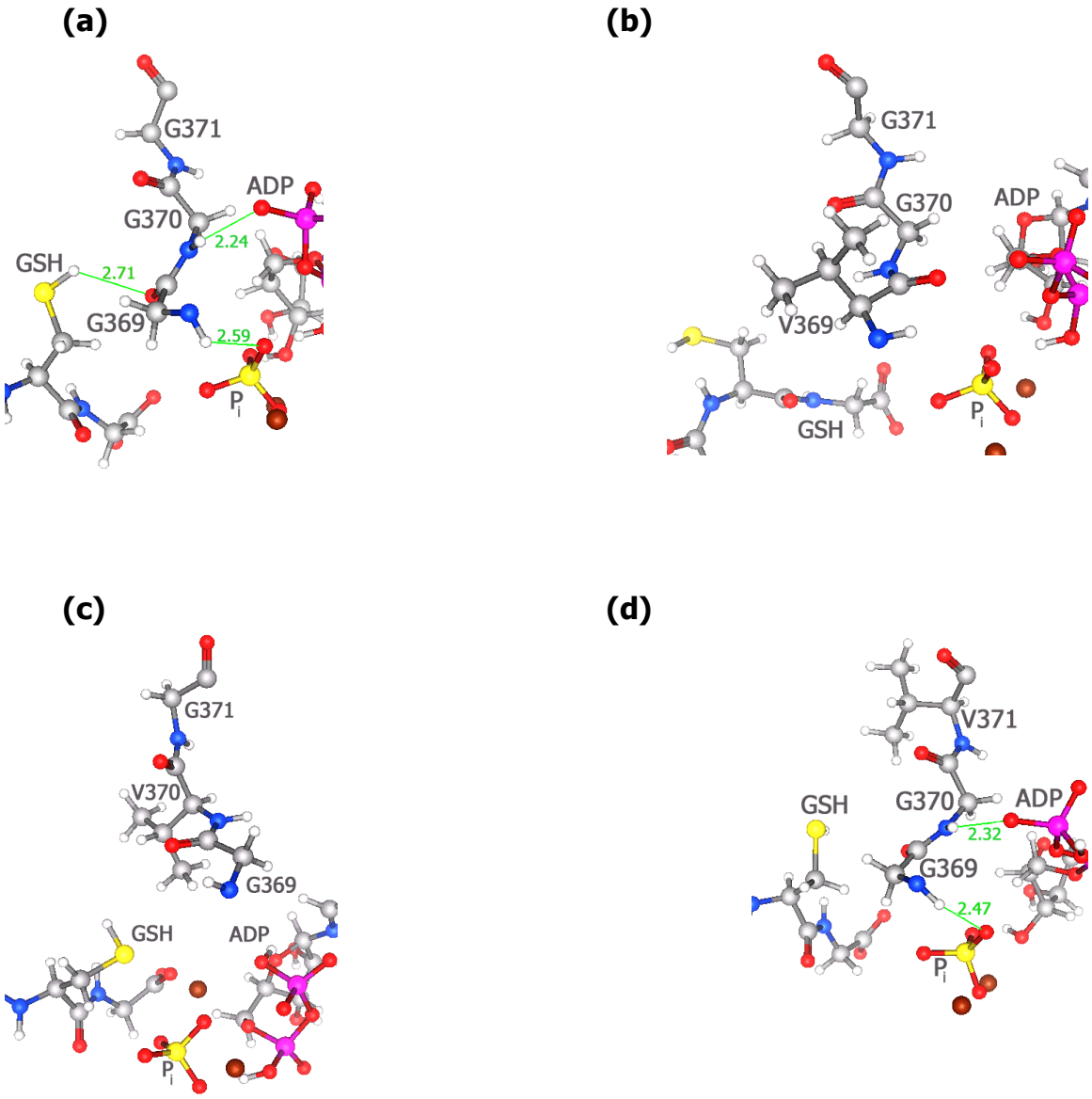


Figure 7.2. Hydrogen bonding between the ligands and G-loop residues in the product complex conformation. (a) wild type; (b) G369V; (c) G370V; (d) G371V.

7.3.2. Ligand Binding in the Product Conformation

Analysis of the structure of the GS wild type product complex (Figure 7.2.a) shows that Gly369, Gly370 and Gly371 form hydrogen bonds with the products (GSH, P_i, ADP). Thus, Gly369 forms two hydrogen bonds: one with inorganic phosphate (P_i; 2.59 Å) and one with the cysteine moiety of GSH (2.71 Å). An oxygen atom on the β-phosphate of ADP (2.24 Å) forms a hydrogen bond with Gly370. In contrast with the reactant complex, Gly371 no longer hydrogen bond with the products in the product complex structure. Interestingly, Gly369 remains in contact with the γ-phosphate in both reactant and product conformations suggesting that Gly369 is involved in phosphate transfer to the γ-glutamylcysteine substrate. Comparison of the wild type product structure with that of the mutant product structures thus provides insight into important interactions of hGS with its ligands, as well as the effects caused by mutations of these key glycine residues.

Analysis of the G369V product structure shows that none of the contacts detected in the wild type product conformation are possible (Figure 7.2.b). Echoing the findings for the reactant complex structure (described above), there is a different orientation of the backbone amide groups of Val369 and Gly370 that obstructs the hydrogen bonding. Analysis of the structure of G370V mutant enzyme in the product conformation also shows the orientations of ADP and P_i are different from those of the wild type. In wild type hGS, Gly369 forms hydrogen bonds with both the P_i and GSH, but for G370V mutant hGS these hydrogen bonds no longer exist. Similarly, where the wild type Gly370 amide group forms a hydrogen bond with an oxygen of the

β–phosphate of ADP (Figure 7.2.a), this bond is not present in the G370V mutant enzyme (Figure 7.2.c). Analogous to the reactant complex findings (above), analysis of the G371V product-enzyme complex shows little change in structure in comparison with the wild type product (Figure 7.2.d). Thus, for the G371V mutant enzyme, Gly369 can still effectively hydrogen bond with the P_i (2.47 Å), although there is no longer any contact with cysteine moiety of GSH. In addition, Gly370 still forms a hydrogen bond (2.32 Å) with the β-phosphate of ADP. The modeling results for both the reactant and product form of wild type and mutant hGS thus suggest that Gly369 and Gly370 have essential roles for the G-loop conformation and ATP binding, while Gly371 has a lesser involvement in ATP binding.

7.3.3. G-loop: A-loop Interaction

In addition to loop-ligand interactions, there are important loop-loop interactions in human GS. This section analyzes the change in these interactions with mutation of the three consecutive backbone glycines (369 to 371). The interaction between the G-loop and A-loop of the wild type hGS occurs only in the reactant conformation (Figure 7.3.a). The carbonyl oxygen of Gly369 forms hydrogen bonds with the amide groups of both Val461 (2.99 Å) and Ala462 (2.44 Å), suggesting that this interaction enables active site closure, which may play a role in hGS catalysis.

Based on substrate orientations and loop positions, the G-loop: A-loop interaction (loop closure), may also be important for stabilizing the acyl-phosphate intermediate and/or transition state for peptide bond formation. Hydrogen bonding patterns show

that loop-loop interaction orients both the γ-phosphate of ATP and the glycine substrate near (~10 Å) the γ-glutamylcysteine substrate. In the reactant conformation, Gly369, Gly370, and Gly371 bind the γ-phosphate (as described above), while Gly460 of the A-loop binds the glycine substrate by forming two hydrogen bonds with distances of 2.05 Å and 3.08 Å. Thus, loop closure and glycine residues in the G-loop, particularly Gly369 and Gly370, have a significant contribution to γ-phosphate transfer from ATP to the γ-glutamylcysteine while the A-loop residues (Gly460, Val461, Ala462) together with Gly369 may facilitate the glycine substrate transfer to the acyl-phosphate intermediate.

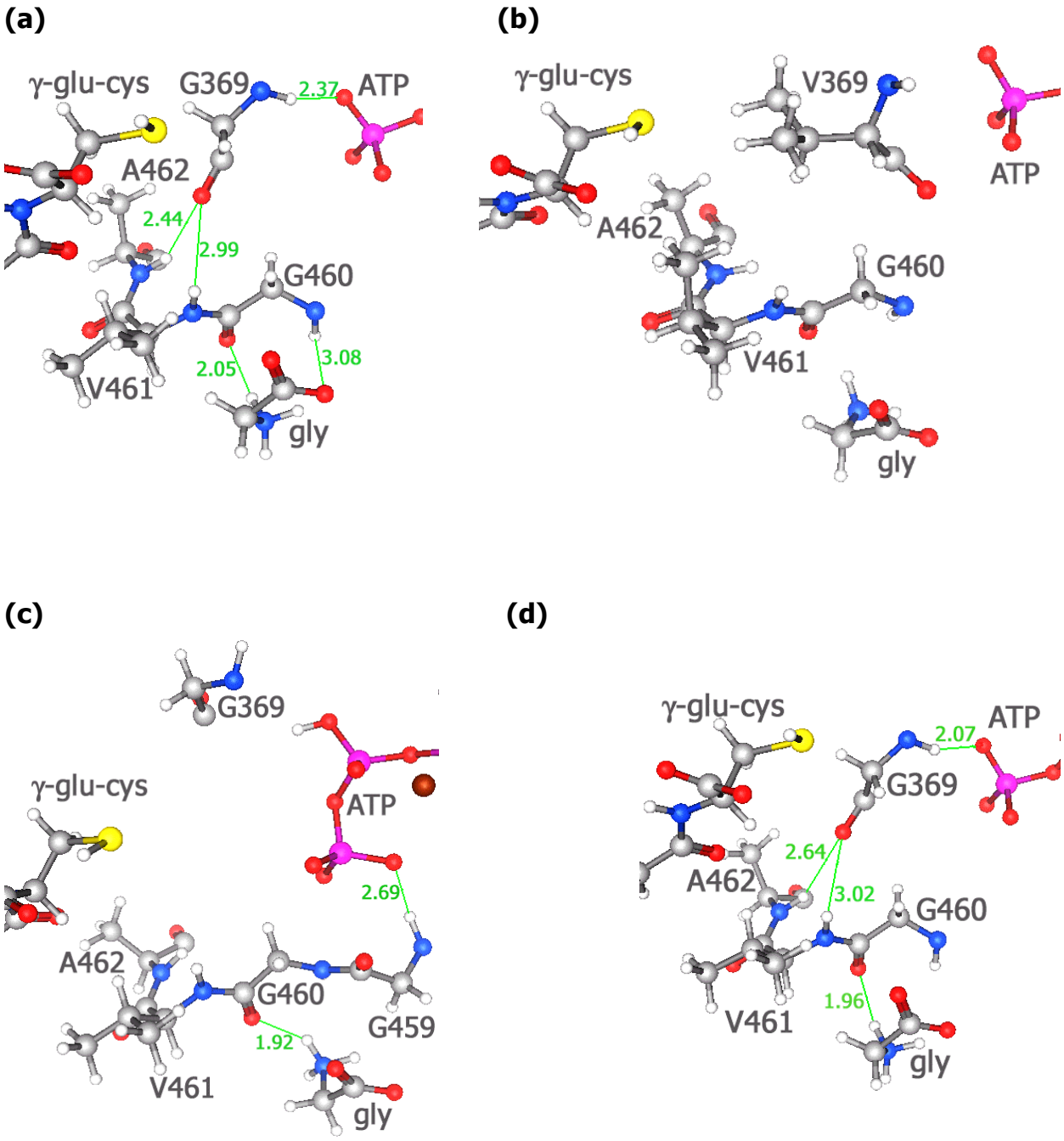


Figure 7.3. G-loop: A-loop interaction and glycine binding in the reactant complex conformation. The Gly370 and Gly371 residues were omitted for clarity. (a) wild type; (b) G369V; (c) G370V; (d) G371V.

The mutation of glycine 369 to valine, a bulky side chain, alters the G-loop backbone orientation of G369V (reactant form) so that loop closure is no longer possible (Figure 7.3.b). Compared to the wild type enzyme, in the G369V mutant hGS structure the carbonyl oxygen of Val369 now points in the opposite direction, which prohibits its hydrogen bonding with Val461 and Ala462. When Gly370 is mutated to valine (Figure 7.3.c), the distance between the Gly369 (G-loop) and A-loop residues increases significantly from *ca.* 3 Å to *ca.* 8 Å between the oxygen of Gly369 and amide groups of Val461 and Ala462, disrupting loop-loop interaction. Structural analysis of the mutation of the glycine residue at 371 to valine shows that there is a similar geometry (Figure 7.3.d) for both ligand binding and G-loop: A-loop interaction as described in Chapter 6 for the wild type enzyme. In G371V mutant hGS, the oxygen atom of Gly369 forms hydrogen bonds with the amide groups of Val461 (3.02 Å) and of Ala462 (2.64 Å), which are of similar length to those calculated in the wild type enzyme. In the same manner, Gly460 hydrogen bonds to the glycine substrate, albeit less effectively because there is only one hydrogen bond of 1.96 Å between the oxygen atom of G460 and the amide group of the glycine substrate in G371V instead of two hydrogen bonds as in wild-type hGS. These findings suggest that G369V and G370V mutations causes poor ligand binding and prevent loop closure in hGS, while G371V mutation has less of a disruptive effect on hGS activity than mutation of the other two G-loop glycines.

7.4. Conclusions

Analyses of local conformational changes and hydrogen bonding variations help explain the experimentally observed decreased enzyme activity and altered binding

affinity of mutant hGS relative to the wild type. The current study provides valuable structural insights on wild type and mutant hGS enzymes regarding the function of the highly conserved Gly369-Gly371 residues of the G-loop. Backbone rotations around the C-C and C-N bonds in protein structures are known to generate more flexible dihedral angles for glycine than other amino acid residues.[100] Mutation of glycine residues in mobile loops to bulkier residues, such as valine, is expected to reduce the loop flexibility and, accordingly, one may thus probe the effect of limited loop flexibility on catalytic efficiency. Both G369V and G370V mutant enzymes have extremely low activities as measured experimentally (Anderson, M. E. and coworkers, Texas Woman's University, personal communication), while G371V is less affected, but still shows a *ca.* 10-fold decrease in k_{cat}. It is gratifying that these experimental results, made subsequent to the computational prediction, are in accord.

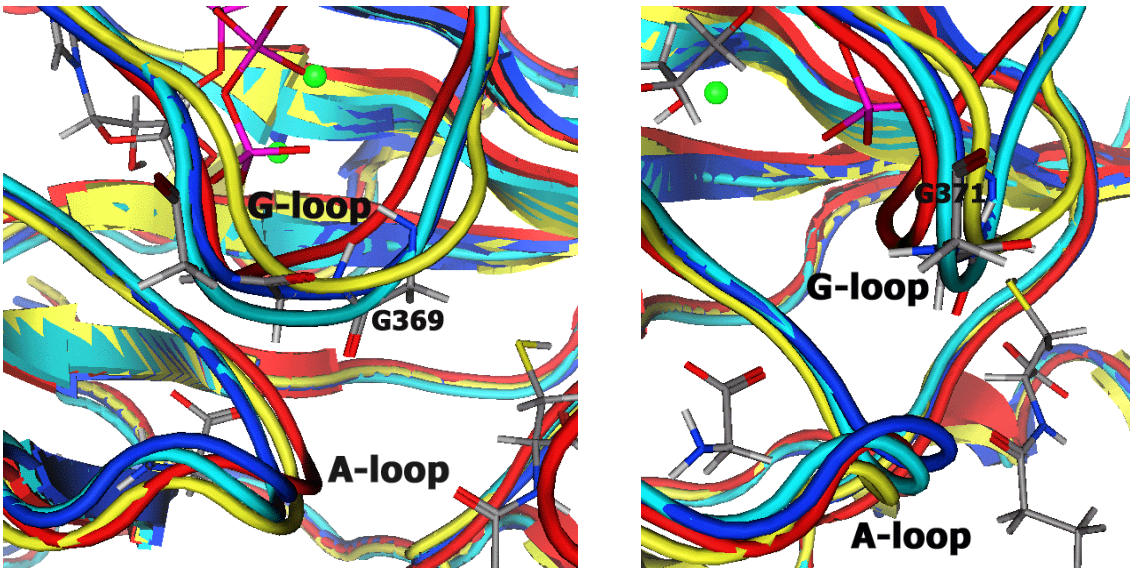


Figure 7.4. Distortion of the G-loop backbone in reactant-complex for the mutant enzymes. Front view (left) and side view (right) show both G and A loops. The wild type is rendered in blue, G369V in red, G370V in yellow, and G371 in cyan.

By probing the effect of single-point mutations (G369V, G370V, and G371V) this work clarifies the role of the three glycine residues in the G-loop of hGS at a molecular level. Computer modeling analyses of the interaction between the G-loop and ligands for each mutant of hGS indicate distinct orientational differences as compared to wild-type hGS vis-à-vis the glycine residues, which affect their hydrogen bonding arrangements. As a result, the binding of γ-phosphate of ATP is altered, generating mutant enzymes with substantially reduced activity. The differences in mutant enzymes (Figure 7.4) are caused by local conformational changes that, consequently, affect the cofactor orientation and probably, transition state stabilization.

The hydrogen bonding pattern of the wild type hGS provides considerable information about the role of these three glycine residues when both reactant and product enzyme complexes are considered. In the reactant conformation, all three glycine residues bind the γ-phosphate of ATP, while in the product structure only two of them are involved: Gly369 binds the cleaved γ-phosphate and Gly370 binds the β-phosphate. Hydrogen bonds between G371 and cofactor are not present in the product-complex enzyme. All these favorable interactions detected in the wild type enzyme suggest that the G369 and G370 are responsible for the γ-phosphate transfer to the γ-glutamyl-cysteine, while G371 has only a minor contribution to the correct orientation of the γ-phosphate. These results suggest that, as observed *in vitro* by Anderson and coworkers (unpublished results), the G371V mutant should have some activity and this single-point mutation should have little affect on the ATP K_m value.

Each glycine mutation to valine in the G-loop reduces the backbone rotation, leading to reduced phosphate binding and a decreased rate for its transfer to the substrate. Based on the trace activity of G369V and G370V mutant enzymes and computer modeling, it is expected that glycine residues, 369 and 370, which directly participate in catalysis by hydrogen bonding to ATP, are necessary for glutathione synthesis while Gly371 is less essential. The three glycine residues are strictly conserved in the eukaryotic GS enzymes. Thus, the present simulations indicate that their role in the loop structure is to participate in ATP binding. Moreover, the glycine conformational variations provide an increased flexibility for the loop motions that may possibly lead to an induced fit for substrates.

Although the glycine-rich loop motif was found in many nucleotide binding enzymes, hGS displays a rare glycine conserved triad Gly-Gly-Gly sequence that links these three residues consecutively. In contrast, the glycine-rich loop sequence in ALAS,[95] DHFR,[96] and the protein kinase family[97] is Gly-X-Gly-X-X-Gly, where X is a variable amino acid. Other phosphate binding loops[101] that interact with the γ-phosphate of ATP have the consensus sequence Gly-X-X-X-X-Gly, including the large loop of *E. coli* GS enzyme. The unusual glycine sequence in the G-loop of hGS greatly reduces the steric repulsions and enhances the flexibility of the backbone in close proximity to the γ-phosphate of ATP. These factors facilitate a favorable orientation of ATP in the binding site, guide the transfer of γ-phosphate to the substrate, and allow the access of the glycine substrate followed by its nucleophilic attack on the acyl-phosphate intermediate. Although the rate-limiting step is still under debate, the glycine

backbone atoms could easily form hydrogen bonds with the γ-phosphate of ATP in the transition state, lowering the activation energy for glutathione biosynthesis.

REFERENCES

1. Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* 1996, *96*, 2239.

2. Fontecave, M.; Pierre, J. L. *Coord. Chem. Rev.* 1998, *170*, 125.

3. Watt, R. K.; Ludden, P. W. *Cell. Mol. Life Sci.* 1999, *56*, 604.

4. McMaster, J. *Annu. Rep. Prog. Chem., Sect. A* 2004, *100*, 611.

5. Darnault, C.; Volbeda, A.; Kim, E. J.; Legrand, P.; Vernede, X.; Lindahl, P. A.; Fontecilla-Camps, J. C. *Nat. Struct. Biol.* 2003, *10*, 271.

6. Hausinger, R. P. *Nat. Struct. Biol.* 2003, *10*, 234.

7. Liao, J. C.; Sun, S.; Chandler, D.; Oster, G. *Eur. Biophys. J.* 2004, *33*, 29.

8. Boyer, P. D. Nobel Lectures, Chemistry 1996-2000; Grenthe, I. Ed.; World Scientific Publishing, Singapore, 2003.

9. Williams, N. H. *J. Am. Chem. Soc.* 2000, *122*, 12023.

10. Zheng, J.; Knighton, D. R.; Ten Eyck, L. F.; Karlsson, R.; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. *Biochemistry* 1993, *32*, 2154.

11. Strugatsky, D.; Gottschalk, K. E.; Goldshleger, R.; Karlish, S. J. D. *Biochemistry* 2005, *49*, 15961.

12. Roothaan, C. C. J. *Rev. Mod. Phys.* 1951, *23*, 69.

13. Schmidt, M. V.; Gordon, M. S. Annu. *Rev. Phys. Chem.* 1998, *49*, 233.

14. Roos, B. O.; Taylor, P. R.; Siegbahn, E. M. *Chem. Phys.* 1980, *48*, 157.

15. Koch, W.; Holthausen, M. C. A Chemist's Guide to Density Functional Theory; Wiley-VCH: Weinheim, Germany, 2000.

16. (a) Becke, A. D. *J. Chem. Phys.* 1993, *98*, 1372. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* 1988, *37*, 785.

17. Jensen, F. Introduction to Computational Chemistry, John Wiley & Sons, Chichester, England, 1999.

18. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* 1995, *17*, 5179.

19. Schlick, T. Molecular Modeling and Simulation – An Interdisciplinary Guide, Springer- Verlag, New York, 2002.

20. Muller, P.; Fruit, C. *Chem. Rev.* 2003, *103*, 2905; Vedernikov, A. N.; Caulton, K. G. *Chem. Comm.* 2004, *2*, 162.

21. Mindiola, D. J.; Hillhouse, G. L. *Chem. Comm.* 2002, 1840.

22. Bart, S. C.; Lobkovsky, E.; Bill, E.; Chirik, P. J. *J. Am. Chem. Soc.* 2006, *128*, 5302.

23. (a) Brown, S. D.; Betley, T. A.; Peters, J. C. *J. Am. Chem. Soc.* 2003, *125*, 322-323. (b) Jenkins, D. M.; Betley, T. A.; Peters, J. C. *J. Am. Chem. Soc.* 2002, *124*, 11238.

24. Tafesh, A. M.; Weiguny, *J. Chem. Rev.* 1996, *96*, 2035.

25. Paul, P. *Coord. Chem. Rev.* 2000, *203*, 269.

26. Moiseev, I. I.; Stromnova, T. A.; Vargaftik, M. N.; Orlova, S. T.; Chernysheva, T. V.; Stolarov, I. P. *Catalysis Today* 1999, *51*, 595.

27. (a) Tsumura, R.; Takagi, U.; Masaki, T.; Ikeda, K. *Jpn. Kokai Tokkyo Koho* 1979, 8 pp. JP 78-67629 19780607. (b) Izumi, Y. *Jpn. Kokai Tokkyo Koho* 1992, 4 pp. JP 90-198013 19900727. (c) Grolig, J.; Rasp, C.; Scharfe, G.; Swodenk, W. *Ger. Offen.* 1979, 18 pp. DE 77-2750282 19771110.

28. (a) Liu, I. O. Y.; Cant, N. W. *J. Catal.* 2005, *230*, 123. (b) Liu, I. O. Y.; Cant, N. W.; Haynes, B. S.; Nelson, P. F. *J. Catal.* 2001, *203*, 487.

29. Kober, E. H.; Martin, R. H.; Raymond, M. A. *U. S. Patent* 1975, 5 pp. US 3884952 19750520.

30. Hamilton, C. W.; Laitar, D. S.; Sadighi, J. P. *Chem. Commun.* 2004, 1628.

31. Lee, S. W.; Trogler, W. C. *Inorg. Chem* 1990, *29*, 1659.

32. Chen, H.; Olmstead, M. M.; Shoner, S. C.; Power, P. P. *J. Chem. Soc. Dalton Trans.* 1992, 451.

33. Gambarotta, S.; Bracci, M.; Floriani, C.; Chiesi-Villa, A.; Guastini, C. *J. Chem. Soc. Dalton Trans.* 1987, 883.

34. Blue, E. D.; Davis, A.; Conner, D.; Gunnoe, T. B.; Boyle, P. D.; White, P. S. *J. Am. Chem. Soc.* 2003, *125*, 9435.

35. Badiei, Y. M.; Krishnaswamy, A.; Melzer, M. M.; Warren, T. H. *J. Am. Chem. Soc.* 2006, *128*, 15056.

36. Frisch, M. J.; Pople, J. A. *et. al*, Gaussian 03, Revision C.02, Gaussian Inc., Wallingford, CT, 2004.

37. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T., Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comp. Chem.* 1993, *14*, 1347.

38. Shoemaker, J. R.; Burggraf, L. W.; Gordon, M. S. *J. Phys. Chem*. A 1999, *103*, 3245.

39. (a) Ponder, J. W.; Richards, F. M. *J. Comp. Chem.* 1987, *8*, 1016. (b) Kundrot, C. E.; Ponder J. W.; Richards, F. M. *J. Comp. Chem.* 1991, *12*, 402.

40. Allinger, N. L.; Yuh, Y. H.; Lii, J. H.; *J. Am. Chem. Soc.* 1989, *111*, 8551.

41. Stevens, W. J.; Basch, H.; Krauss, M.; Jasien, P. G. *Can. J. Chem.* 1992, *70*, 610.

42. Borden, W. T.; Gritsan, N. P.; Hadad, C. M.; Karney, W. L.; Kemnitz, C. R.; Platz, M. S. *Acc. Chem. Res.* 2000, *33*, 765.

43. Warren, T. H. 231st ACS National Meeting, Atlanta, GA, March 26-30, 2006, INOR-55.

44. Blue, E. D.; Gunnoe, T. B.; Petersen, J. L.; Boyle, P. D. 230th ACS National Meeting, Washington, DC, Aug. 28-Sept. 1, 2005, INOR-260.

45. Holland, P. L.; 231st ACS National Meeting, Atlanta, GA, March 26-30, 2006, INOR-404.

46. (a) Cundari, T. R.; Gordon, M. S. *J. Am. Chem. Soc.* 1991, *113*, 5231. (b) Taylor, T. E.; Hall, M. B. *J. Am. Chem. Soc.* 1984, *106*, 1576.

47. Carter, E. A.; Goddard, W. A., III. *J. Phys. Chem.* 1988, *92*, 2109.

48. (a) Badiei, Y. M.; Warren, T. H. *J. Organomet. Chem.* 2005, *690*, 5989. (b) Dai, X.; Warren, T. H. *J. Am. Chem. Soc.* 2004, *126*, 10085. (c) Kogut, E.; Wiencko, H. L.; Zhang, L.; Cordeau, D. E.; Warren, T. H. *J. Am. Chem. Soc.* 2005, *127*, 11248.

49. Cundari, T. R. *J. Am. Chem. Soc.* 1992, *114*, 7879.

50. Cundari, T. R.; Russo, M. *J. Chem. Info. Comput. Sci.* 2001, *41*, 281.

51. Sydora, O. L.; Kuiper, D. S.; Wolczanski, P. T.; Lobkovsky, E. B.; Dinescu, A.; Cundari, T. R. Inorg. Chem. 2005, 45, 2008.

52. Murray, B. D.; Hope, H.; Power, P. P. *J. Am. Chem. Soc.* 1985, *107,* 169.

53. Edema, J. J. H.; Gamabarotta, S.; Speck, A. L. *Inorg. Chem.* 1989, *28*, 811.

54. Ruppa, K. B. P.; Feghali, K.; Kovacs, I.; Aparna, K.; G
Bensimon, C. *J. Chem. Soc., Dalton Trans.* 1998, *10*, 1595.

55. Hirao, K. *Chem. Phys. Lett.* 1993, *201*, 59

71. Bains, J. S.; Shaw, C. A. *Brain Res. Brain Res. Rev.* 1997, *25*, 335.

72. Meister, A. Glutathione synthesis; Boyer, P. D. Ed.; The Enzymes, third ed. vol.10, Academic Press, New York, 1974.

73. Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. *J. Mol. Biol.* 1995, *247*, 536.

74. Galperin, M. Y.; Koonin, E. V. *Protein Sci.* 1997, *6*, 2639.

75. Heinemann, U.; Hahn, M. *Prog. Biophys. Molec. Biol.* 1995, *64*, 121.

76. Dinescu, A.; Cundari, T. R.; Bhansali, V. S.; Luo, J. L.; Anderson, M. E. *J. Biol. Chem.* 2004, *279,* 22412.

77. Tanaka, T.; Yamaguchi, H.; Kato, H.; Nishioka, T.; Katsube, Y.; Oda, J. *Biochemistry* 1993, *32*, 12398.

78. Kato, H.; Tanaka, T.; Yamaguchi, H.; Hara, T.; Nishioka, T.; Katsube, Y.; Oda, J. *Biochemistry* 1994, *33*, 4995.

79. Polekhina, G.; Board, P. C.; Gali, R.; Rossjohn, J.; Parker, M. V. *EMBO J.* 1999, *12,* 3204.

80. Gogos, A.; Shapiro, L. *Structure* 2002, *10*, 1669.

81. Matsuda, K.; Mizuguchi, K.; Nishioka, T.; Kato, H.; Go, N.; Oda, J. *Protein Eng.* 1996, *9*, 1083.

82. Hara, T.; Kato, H.; Katsube, Y.; Oda, J. *Biochemistry* 1996, *35*, 11967.

83. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* 2000, *28*, 235.

84. Molecular Operating Environment (MOE), version 2003.02, 2003, Chemical Computing Group, Inc., Montreal, Quebec, Canada.

85. Gasteiger, J.; Marsili, M. *Tetrahedron* 1980, *36*, 3219.

86. Dinescu, A. Modeling wild type and mutant glutathione synthetase, Thesis (M.S.), University of North Texas, Denton, Texas, 2004.

87. DeepView/Swiss-PdbViewer, version 3.7 (SP4), 2003, Swiss Institute of Bioinformatics, Lausanne, Switzerland.

88. O'Donovan, C.; Martin, M. J.; Gattiker, A.; Gasteiger, E.; Bairoch, A.; Apweiler, R. *Brief. Bioinform.* 2002, *3*, 275.

89. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* 1994, *22*, 4673.

90. Szustakowski, J. D.; Weng, Z. *Proteins* 2000, *38*, 428.

91. Tanaka, T.; Kato, H.; Nishioka, T.; and Oda, J. *Biochemistry* 1992, *31*, 2259.

92. Pan, Q. W.; Tanase, S.; Fukumoto, Y.; Nagashima, F.; Rhees, S.; Rogers, P. H.; Amone, A.; Morinof, Y. *J. Biol. Chem.* 1993, *268*, 24758.

93. Okoniewska, M.; Tanaka, T.; Yada, R. Y. *Biochem. J.* 2000, *349*, 169.

94. Pompliano, D. L.; Peyman, A.; Knowles, J. R. *Biochemistry* 1990, *29,* 3186.

95. Gong, J.; Kay, C. J.; Barber, M. J.; Ferreira, G. C. *Biochemistry* 1996*, 35,* 14109.

96. Miller, G. P.; Benkovic, S. J. *Biochemistry* 1998, *37,* 6327.

97. Hirai, T. J.; Tsigelny, I.; Adams, J. A. *Biochemistry* 2000, *39,* 13276.

98. Grant, B. D.; Hemmer, W.; Tsigelny, I.; Adams, J. A.; Taylor, S. S. *Biochemistry* 1998*, 37,* 7708.

99. Hemmer, W.; McGlone, M.; Tsigelny, I.; Taylor, S. S. *J. Biol. Chem.* 1997, *272*, 16946.

100. Bright, J. N.; Sansom M. S. P. *J. Phys. Chem. B* 2003, *107*, 627.

101. Saraste, M.; Sibbald, P. R.; Wittinghofer, A. *Trends Biochem. Sci.* 1990*, 15*, 430.